

Portland State University
PDXScholar

Dissertations and Theses

Dissertations and Theses

Winter 4-4-2017

Adaptive Evolution under Favorable and Unfavorable Population Genetic Conditions in *Caenorhabditis* *elegans* Nematodes

Stephen Fuller Christy
Portland State University

Follow this and additional works at: https://pdxscholar.library.pdx.edu/open_access_etds



Part of the [Biology Commons](#)

Let us know how access to this document benefits you.

Recommended Citation

Christy, Stephen Fuller, "Adaptive Evolution under Favorable and Unfavorable Population Genetic Conditions in *Caenorhabditis elegans* Nematodes" (2017). *Dissertations and Theses*. Paper 3534.
<https://doi.org/10.15760/etd.5420>

This Thesis is brought to you for free and open access. It has been accepted for inclusion in Dissertations and Theses by an authorized administrator of PDXScholar. Please contact us if we can make this document more accessible: pdxscholar@pdx.edu.

Adaptive Evolution under Favorable and Unfavorable Population Genetic Conditions in

Caenorhabditis elegans Nematodes

by

Stephen Fuller Christy

A thesis submitted in partial fulfillment of the
requirements for the degree of

Master of Science
in
Biology

Thesis Committee:
Suzanne Estes, Chair
Jason Podrabsky
Todd Rosenstiel

Portland State University
2017

ABSTRACT

Mutation is a fundamental process that drives evolutionary change; however, most new mutations are deleterious for organismal fitness and can readily propagate within populations under a broad range of conditions. Mutational processes able to counteract deleterious mutation accumulation include: 1) reversion mutation back to wildtype, 2) acquisition of generally beneficial mutations, and 3) compensatory mutations that specifically mitigate the effects of previously-acquired deleterious mutations through epistasis. The potential for any of these mutation types alters our expectations for the impact of deleterious mutation in populations, but since the fitness effects of individual mutations are rarely characterized, the relative importance of beneficial and compensatory epistatic mutations is unknown. In this thesis, I characterized the nuclear mutations that arose in a previous mutation accumulation (MA) experiment using *Caenorhabditis elegans* nematodes, in which mutations were allowed to accumulate under extreme drift conditions in replicate, independently evolving lines initiated from a low-fitness mitochondrial electron transport chain (ETC) mutant, *gas-1*. In contrast to the results of typical MA experiments, *gas-1* MA lines improved fitness slightly compared to their ancestor. Here, I find that the *gas-1* MA lines demonstrate little increase in among-line variance and that the *gas-1* MA nuclear mutations are more narrowly functionally defined than wildtype MA nuclear mutations. When combined with evidence for zygotic or post zygotic selection these data suggest that selection—both purifying and positive—can be an extremely powerful force even in conditions of extreme genetic drift. Furthermore, functional characterization of a four-mutation set isolated from one of the

gas-1 MA lines on *gas-1* and wildtype backgrounds shows fitness improvements on both backgrounds. This beneficial four-mutation set is associated with a decrease in steady-state endogenous ROS on the *gas-1* background while exhibiting no effect on wildtype. I also find that steady-state ATP levels associated with the beneficial four-mutation set decreased compared to wildtype suggesting that fermentation may be metabolic strategy to cope with increase oxidative stress. These findings suggest that we can detect and characterize specific genetic changes that lead to a partial recovery of fitness and phenotype in a low-fitness ETC-deficient mutant strain of *C. elegans*. I extended my thesis to include analyses of fitness and phenotype of 24 replicate lineages of the *gas-1* ETC mutant evolved in large population ($n = 1000$) sizes for 60 generations—conditions optimal for selection and fitness recovery (RC). I find that two distinct *gas-1* RC fitness groups emerged: one group with significantly higher average fitness than the ancestor and containing two lines that exceeded wildtype fitness levels, and another group with more modest and non-significant fitness gains. Interestingly, many lines in the first group generated appreciable numbers of males during experimental evolution—consistent with evolution of outcrossing either accompanying or driving rapid fitness recovery. Bioinformatic functional analyses of the nuclear mutations that arose in the *gas-1* RC lines show the availability of potentially more paths to fitness recovery for large than small populations. Combined, these data allow us to identify patterns in selection and drift in *gas-1* recovery under MA and RC (recovery) conditions. My research advances our understanding of the genetic bases of adaptive evolution under extremely unfavorable population genetic conditions and how ETC dysfunction affects evolutionary dynamics.

ACKNOWLEDGEMENTS

I would like to thank Dr. Suzanne Estes for all her support and assistance in this project; Drs. Jason Podrabsky and Todd Rosenstiel for their insight and advice serving on my thesis committee; Drs. Dee Denver and Riana Wernick, and Dana Howe at Oregon State University for being wonderful collaborators; Drs. Charles Baer and Levi Morran for help with worms; Jennifer Sullins for her amazing work in the lab; undergrads Alex Basler, Joseph Ramirez, Ashleigh Mustain, Lauren Muñoz-Tremblay, and Griselda Velasco for all their tireless help with projects; Michael Lue for helping me onboard; and finally, my friends and family for their endless encouragement.

TABLE OF CONTENTS

ABSTRACT	i
ACKNOWLEDGEMENTS	iii
LIST OF TABLES	v
LIST OF FIGURES	vi
LIST OF ABBREVIATIONS	vii
CHAPTER 1: INTRODUCTION	1
CHAPTER 2: PHENOTYPIC RECOVERY UNDER UNFAVORABLE POPULATION GENETIC CONDITIONS	16
Background.....	16
Methods.....	21
Results	30
Discussion	45
CHAPTER 3: FITNESS RECOVERY IN LARGE POPULATIONS	54
Background.....	54
Methods.....	58
Results	65
Discussion	76
CHAPTER 4: CONCLUSION	81
REFERENCES.....	85
APPENDIX A: GO term data for <i>gas-1</i> MA lines.....	93
APPENDIX B: GO term data for N2 MA lines	95
APPENDIX C: GO term data for <i>gas-1</i> centric interactome	99
APPENDIX D: GO term data for <i>gas-1</i> RC lines	102

LIST OF TABLES

Table 2.1. Summary statistics for life-history traits.....	35
Table 2.2. Daily offspring production of N2, <i>gas-1</i> G0, and <i>gas-1</i> MA lines.....	36
Table 2.3. Variance components for N2 and <i>gas-1</i> G0 strains.....	37
Table 2.4. Comparisons of daily progeny production of N2, N2 ^{MA431} , <i>gas-1</i> G0, and <i>gas-1</i> ^{MA431} ...	43
Table 3.1. One-way ANOVAs for life history traits for N2, <i>gas-1</i> G0, <i>gas-1</i> RC reported as means or variances \pm S.E.M.....	68
Table 3.2. Means for life history traits for N2, <i>gas-1</i> G0, and <i>gas-1</i> RC.....	69
Table 3.3. Summary statistics for daily reproduction for <i>gas-1</i> G0 and N2.....	70
Table 3.4. Two-way mixed-model ANOVA for life history traits in <i>gas-1</i> RC and <i>gas-1</i> G0. Statistical analyses are detailed in methods.....	72

LIST OF FIGURES

Figure 1.1. Fitness landscapes and the distribution of fitness effects.	5
Figure 1.2. Schematic of mitochondrial ETC supercomplex.	11
Figure 2.1. Reproductive schedules.....	33
Figure 2.2. Frequency histogram of <i>gas-1</i> MA relative fitness.	33
Figure 2.3. Proportions of GO slim categories significantly enriched among the five sequenced <i>gas-1</i> MA lines (A) and five N2 MA lines (B).	39
Figure 2.4. Relative fitnesses for MA431 SNP-containing lines and ancestral controls.	42
Figure 2.5. Reproductive schedule for MA431 SNP-containing lines and ancestral controls.....	42
Figure 2.6. Age-specific survival for MA431 SNP-containing lines and ancestral controls.....	44
Figure 2.7. Steady-state ROS levels for MA431 SNP-containing lines and ancestral controls.....	44
Figure 2.8. Steady-state ATP levels for MA431 SNP-containing lines and ancestral controls.	45
Figure 3.1. Reproductive schedules of individual <i>gas-1</i> RC lines, <i>gas-1</i> G0, and wildtype N2.	67
Figure 3.2. Reproductive schedules of <i>gas-1</i> RC, <i>gas-1</i> G0, and wildtype N2.	67
Figure 3.3. Frequency histogram of the 24 <i>gas-1</i> RC lines' fitness (<i>w</i>) relative to <i>gas-1</i> G0 (equal to 1.0).	71
Figure 3.4. Correlation of relative fitness with competitive fitness.....	73
Figure 3.5. Proportions of GO slim categories significantly enriched among the <i>gas-1</i> RC.	76

LIST OF ABBREVIATIONS

ATP – Adenosine triphosphate
DFE – Distribution of fitness effects
EMS – Ethyl methanesulfonate
ETC – Electron transport chain
GO – Gene ontology
MA – Mutation accumulation
mtDNA – Mitochondrial DNA
nDNA – Nuclear DNA
 N_e – Effective population size
RC – Recovery (e.g. recovery line)
ROS – Reactive oxygen species
SNP – Single nucleotide polymorphism
 W – Absolute fitness
 χ^2 – Chi-square
 ω – Relative fitness

CHAPTER 1: INTRODUCTION

Background. New mutation is the ultimate source of genetic variation in populations. Most newly arising mutations are deleterious for organismal fitness, yet can readily propagate within populations under a broad range of conditions (Crow and Kimura 1970). The accumulation of deleterious mutations can have profound implications for the survival of small populations (Lynch et al. 1990; Charlesworth 1993; Lande 1994; Lynch, Conery, et al. 1995; Whitlock 2000; Estes et al. 2004). The threat of small population survival is confounded by low genetic diversity in these populations, who must depend on new mutation rather than recombination to generate favorable alleles (Soulé 1987). However, the fact that populations persist at high levels of fitness under the constant pressure of deleterious mutations suggests that mutational processes can ameliorate the phenotypic decline caused by acquired deleterious mutations (Estes 2003; Estes, Phillips, et al. 2011).

Three mutational processes able to counteract deleterious mutation accumulation include acquisition of: 1) reversion mutations that restore wildtype sequences, 2) generally beneficial mutations that improve organismal fitness irrespective of genetic background and 2) compensatory mutations that specifically mitigate the effects of previously-acquired deleterious mutations through epistasis (Kimura 1985). Unlike generally beneficial mutations, compensatory mutations are beneficial *only* within the context of a particular deleterious mutation and are, by themselves, either neutral or mildly deleterious. The potential for either beneficial or compensatory mutations alters

our expectations for the impact of deleterious mutation in populations (Whitlock and Otto 1999; Orr 2000; Poon and Otto 2000; Lanfear et al. 2014), but since the fitness effects of individual mutations are rarely characterized, the relative importance of either mutation type is unknown. Furthermore, the extent to which beneficial or compensatory mutations are capable of rescuing fitness in small populations with high extinction risk is understudied.

The earliest evolutionary theories treated mutations as deleterious or beneficial regardless of their current effect on fitness (Haldane 1937; Poon and Otto 2000). However, the idea that mutations were context-dependent, where a mutation with a specific phenotypic effect may be deleterious in one genome and advantageous in another, was eventually proposed (Wright 1968). This context-dependence is in part due to epistatic gene-gene interactions, where the effects of one gene are altered or masked by the presence of a modifier gene (Butcher 1995; Phillips et al. 2000; Phillips 2008). An example of this is seen in complex disease traits like Alzheimer's disease in which there are estimated to be over 100 epistatic interactions where certain genes either promote the development of disease or confer protection (Lobo 2008). This context-dependence, for which compensatory mutations are a prime example, introduces variability in the effect on fitness of newly arising mutations in populations (Poon and Otto 2000).

Compensatory mutational effects may be common since individual mutations are likely to have different selective effects depending on their molecular genetic context (Zhang and Watson 2008; Lanfear 2014). An example of differential selective effects in compensatory mutation is seen in antibiotic and antiviral drug resistance (Levin et al

2000). Fixation of drug resistance mutations comes with a cost of fitness decline in microbial (MacLean et al. 2010) and viral populations (Nijhuis et al. 1999). These populations frequently recover pre-drug-resistance fitness levels through fixation of compensatory mutations, which in the case of the Human Immunodeficiency Virus, results in pre-drug-therapy viral loads. The evolutionary mechanisms of compensatory mutation have been well studied in bacteriophage (Poon and Chao 2005) and bacterial laboratory evolution settings; results of such experiments show rapid—but sometimes incomplete—fitness recovery and a high incidence of parallel mutation fixation in independent experimental lineages (Rokyta et al. 2002; Colegrave and Buckling 2005), although their trajectories to fixation can be complex (Meer et al. 2010). Microbial studies have yielded numerous other insights into features of compensatory adaptation including information on the functional genetic bases of new compensatory mutations (Harcombe et al. 2009; Poon and Chao 2005) and the limitations on available evolutionary trajectories (Poelwijk et al. 2007). However, compared to microbes, multicellular eukaryotes are more complex, and this complexity reduces the probability that a mutation will be adaptive (Orr 2000). Parallel experiments with multicellular eukaryotes are needed to assess the generality of these results for more genetically and phenotypically complex organisms (Halligan and Keightley 2009).

The effects of new mutation can also depend on the starting fitness of the population in which they arise (Orr et al. 2003). Using Fisher's adaptive landscape as a heuristic, we see that when populations move within the fitness landscape, their selection coefficients for mutations change, which ultimately affects their substitution rates

(Lanfear et al. 2014). For instance, populations that exist in fitness valleys will experience an altered distribution of fitness effects (DFE) such that mutations that would be mildly deleterious in a high-fitness population become effectively neutral, mutations that would be deleterious become lethal, and mutations that would be mildly advantageous become strongly so (Figure 1.1). This altered DFE across the fitness landscape is a consequence of both the context dependency of mutations and the population's effective size (N_e) and makes ascertaining their organismal- and population-level consequences difficult. Experiments with well-understood model systems are necessary to study the effects of new mutation in an evolutionary framework.

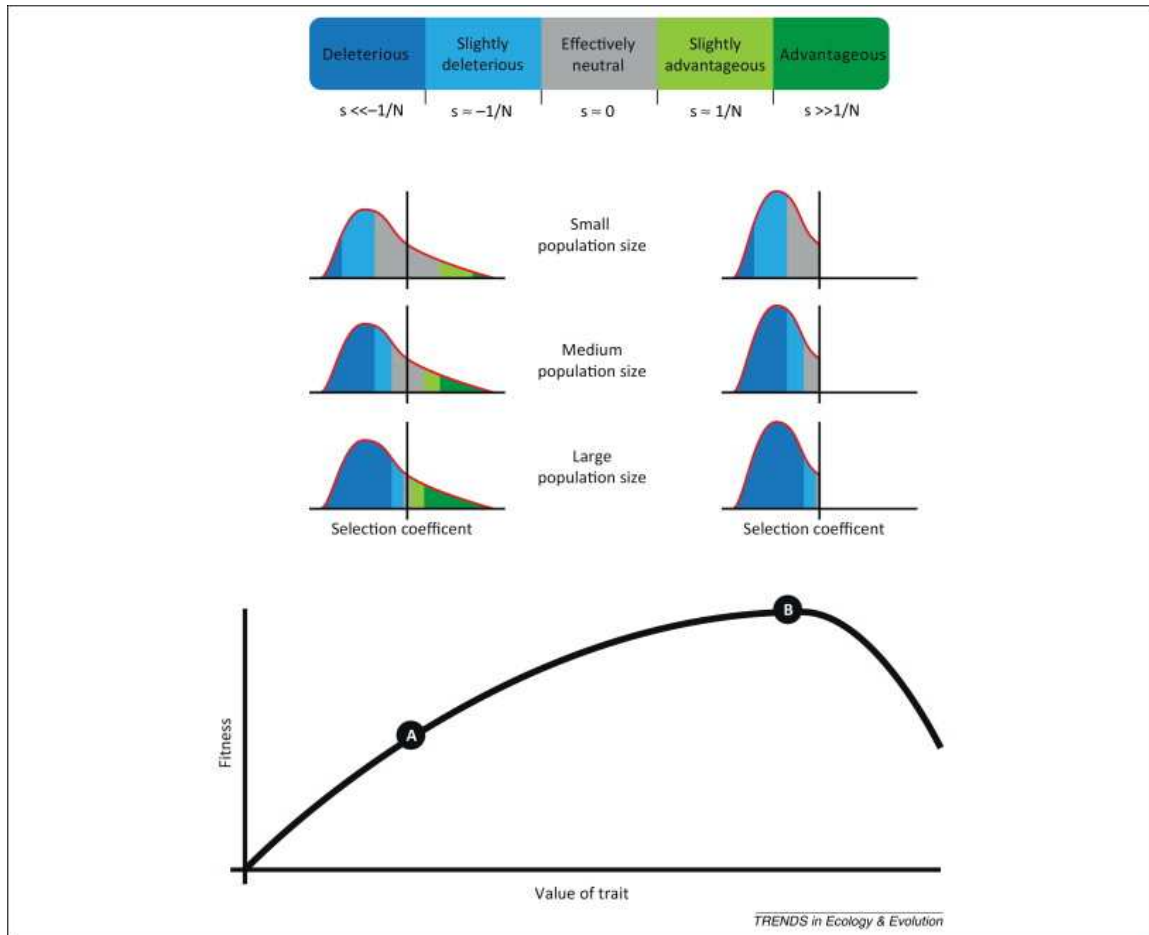


Figure 1.1. Fitness landscapes and the distribution of fitness effects. Hypothetical distributions of fitness effects for populations of different sizes at different positions on a simple fitness landscape. A population far from the optimum (**A**) has a certain proportion of mutations that confer increases in fitness. However, a population at the hypothetical optimum of the landscape (**B**) cannot increase its fitness, so all mutations are deleterious. In both cases, the proportion of mutations that fall into different categories (Box 2, main text) changes depending on the effective population size. Note that, for simplicity, we have drawn a fitness landscape that varies along a single dimension, but the distributions we have drawn are more similar to those that would come from higher-dimensional fitness landscapes. Furthermore, it is unlikely that any natural population sits at the precise optimum of any fitness landscape. The selection coefficients are shown on natural scales, not log-transformed scales. Imaged and text copied from (Lanfear et al. 2014).

Exploring evolution in real time with experimental evolution studies. Many aspects of evolutionary theory are difficult to test and must be inferred (Buckling et al. 2009). Phylogenetic data can tell you what has happened in the past; however, it is difficult to test evolutionary theories in an uncontrolled environment. Experimental evolution studies offer a framework to test evolutionary models in real time (Kawecki et al. 2012).

Experimental evolution is often conducted in the laboratory by evolving replicate lineages of populations for many generations, and then comparing ancestral and evolved lineages. Experimental evolution allows researchers to study diverse evolutionary hypotheses about populations in response to experimenter-imposed environmental, genetic, or demographic conditions (Kawecki et al. 2012). Experimental evolution is therefore a powerful tool to test evolutionary theories—such as how compensatory and generally beneficial mutations alter extant models relating mutation to fitness and mutation accumulation. The advent of next generation sequencing allows us to study in detail the molecular genetic underpinnings of phenotypic and fitness changes over evolutionary time.

The most straightforward experimental evolution studies are mutation-accumulation (MA) experiments wherein mutations are allowed to accumulate in replicate lineages evolving in population sizes of one (asexual or self-fertilizing organisms) or two individuals (sexual species) creating a scenario for extreme genetic drift (Halligan and Keightley 2009). Since most new mutations are mildly deleterious for organismal fitness (Crow and Kimura 1970), MA experiments create population genetic environments where non-lethal deleterious mutations can readily propagate. The effective population size (N_e), the number of individuals who produce offspring for the next generation, plays an important role in determining the evolutionary fate of mutations of mild fitness and phenotypic effect (Kimura 1983; Ohta 1992). As N_e decreases, mutations of mild effect become effectively neutral; random sampling (genetic drift) dictates the evolutionary trajectory of these mutations (Ohta 1992). As a result, MA

lines experience substantially more deleterious mutation accumulation and fitness decline than do larger populations (Estes et al. 2004; Taylor et al. 2002). MA experiments allow researchers to directly estimate mutation rates by comparing whole-genome sequences of MA populations to that of their ancestor. It is also possible to infer the distribution of mutational fitness effects by comparing fitness data from MA and ancestral populations.

In contrast to MA experiments, it is also possible to study the effect of new mutation in large populations. Populations with a large N_e experience more efficient selection and, as a result, are able to effectively purge deleterious mutations. Large populations are not inherently immune from deleterious mutation accumulation, however. The ability of natural selection to purge deleterious mutation can be overwhelmed if the mutation rate is sufficiently high. Larger populations also benefit from fixation of rare favorable mutations owing to the increased number of genomes available for mutation to occur and more efficient selection than for small populations (Crow and Kimura 1970). Previous experiments (Estes and Lynch 2003; Denver et al. 2010; Estes et al. 2011b) have documented the ability of MA lines to recover fitness in large populations. These studies found that large population treatments can restore, at least partially, ancestral levels of fitness and that rapid adaptation is influenced by the fixation of new mutations despite the high probability that such mutations would be lost to drift. These findings suggest that mutations with strong selective effects are able to sweep to fixation even though mutations with such high selective effects are rare.

***Caenorhabditis elegans* in experimental evolution.** The microscopic *Caenorhabditis elegans* nematode is an excellent eukaryotic model for studying molecular-genetic and phenotypic evolution (Gray and Cutter 2014). There are several aspects of *C. elegans*' biology that give them utility in experimental evolution studies. *C. elegans* nematodes have a short generation time of three to four days (Wood 1988; Riddle et al. 1997), which allows for many generations of evolution to occur in a relatively short amount of time. In addition, *C. elegans* is androdioecious, with populations consisting primarily of self-fertile hermaphrodites and rare males (Wood 1988; Riddle et al. 1997); this allows multiple lineages to be generated in the laboratory with an N_e as small as one individual. Wildtype laboratory strains of *C. elegans*, like the canonical Bristol N2 strain (Brenner 1974), progress across four larval stages to adulthood in two to three days and lay roughly 300 eggs over a period of about a week (Wood 1988; Riddle et al. 1997). Wildtype *C. elegans* typically live about two weeks; however, survival and lifespan can be heavily influenced by exogenous and endogenous stressors. It is easy to assess lifespan and survival by monitoring a worms response to gentle prodding with a platinum or titanium metal wire (Stiernagle 2006). Since *C. elegans* can be cryogenically preserved (Stiernagle 2006), ancestral and evolved populations can be directly compared in phenotyping and genotyping experiments. It is also relatively easy to cross *C. elegans*—using males and hermaphrodites—to generate isogenic strains with different mutations that confer experimentally useful phenotypes (Hope 1999). Additionally, *C. elegans* nematodes are highly tractable in laboratory environments and a diverse array of experiments can be performed on them (Hope 1999; Riddle et al. 1997; Hope 1999; Gray and Cutter 2014).

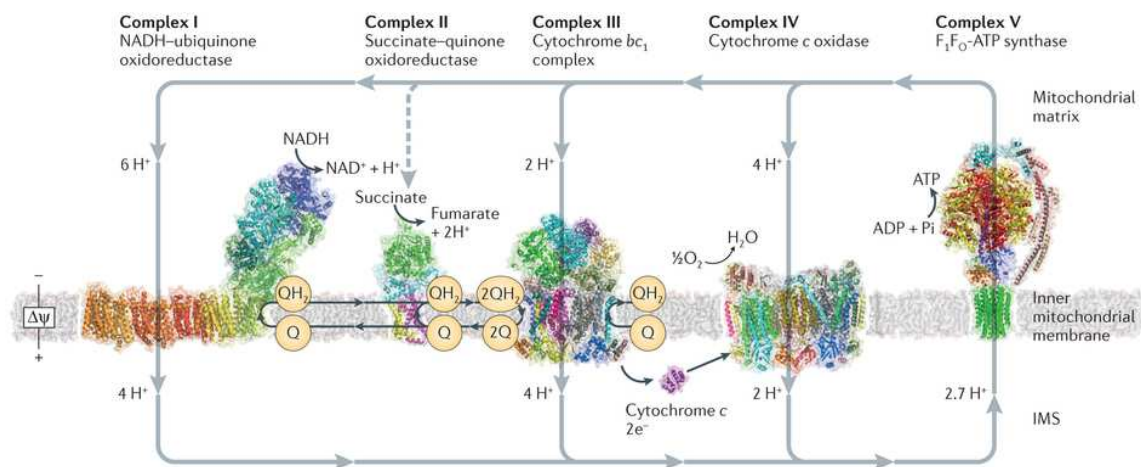
And because the nematodes are nearly transparent, they are well suited to *in vivo* microscopy studies.

A wealth of knowledge has been gained from nearly four decades of study on *C. elegans*, which has led to the creation of numerous research tools. *C. elegans* was the first metazoan genome to be sequenced in 1998 (Equence et al. 1998). In addition, numerous genome-wide RNA interference screens have been performed on *C. elegans* to determine the function of- and interaction between genes (Cipriani and Piano 2011). The computational data from these scientific discoveries in *C. elegans* has been integrated with data from other organisms unlocking numerous high-throughput techniques for identifying and characterizing the functions and interactions of gene products (Zhong and Sternberg 2006).

Mitochondrial electron transport chain (ETC). The mitochondrial electron transport chain (ETC) offers a compelling framework for testing evolutionary hypotheses. Mitochondria are essential eukaryotic organelles that generate the majority of cellular bioenergy in the form of ATP. The ETC is a supercomplex of five enzymatic proteins composed of subunits encoded by both the nuclear and mitochondrial genomes (Figure 1.2). Together, three of the enzymes—complexes I, II, and IV—pump protons across the mitochondrial intermembrane creating a proton motive force, which drives the production of ATP. Because this process, which is vital for energy metabolism in all eukaryotic life, is highly conserved, the ETC is an ideal system in which to study the impact of mutation. ETC integrity is presumably under constant threat as many features of mitochondrial biology—absence of recombination, maternal inheritance, and proximity to damaging

oxygen radicals—have the potential to promote faster rates of, presumably deleterious, evolution in mitochondrial DNA (mtDNA) (Ballard and Whitlock 2004).

In addition, since proper ETC functioning relies on the maintenance of favorable mitonuclear epistatic interactions (Blier 2001; Dowling et al. 2008), it is also an excellent model for addressing epistatic gene interactions and their role in phenotypic evolution. Despite the high mutation rates measured for mtDNA genomes (Denver etc. refs), there is extensive evidence supporting its effective coevolution with the nuclear genome. Mitonuclear mismatch studies, where cytoplasmic and nuclear components are mixed between separate lineages, often result in hybrid breakdown resulting from ETC deficiencies and fitness costs (Nagao et al. 1998; Ballard and Whitlock 2004; Ellison 2006). Interestingly, a recent phylogenetic study (Osada and Akashi 2011) revealed how mitonuclear matching might be maintained by the nuclear genome, which acts as the primary compensator for mitochondrial-encoded ETC dysfunction. Discerning the relative roles of deleterious and compensatory mutation in ETC gene coevolution is an ideal problem for applying a laboratory experimental evolution approach.



Nature Reviews | Molecular Cell Biology

Figure 1.2. Schematic of mitochondrial ETC supercomplex. Protons are pumped from the mitochondrial matrix side of the inner mitochondrial membrane (IMM) to the intermembrane space (IMS) side. QH₂ = ubiquinol, Q = ubiquinone, and $\Delta\Psi$ membrane potential. The PDB file for the ATP synthase was provided by J. E. Walker, and the ETC image was prepared by G. Minhas, Medical Research Council, Mitochondrial Biology Unit, Cambridge, UK. Image and text copied from (Sazanov 2015).

Mitochondrial dysfunction of the *gas-1 (fc21)* phenotype. Here, I report on the capacity of new mutation to resuscitate the unhealthy phenotype of an ETC-deficient *gas-1 (fc21)* *C. elegans* strain compared to the wildtype Bristol N2 strain in real-time under conditions of extreme genetic drift (i.e., mutation accumulation) and efficient selection (i.e., large N_e) using an ETC-deficient *gas-1 (fc21)* *C. elegans* nematode strain compared to the wildtype Bristol N2 strain. The *gas-1* gene, so named for its **G**eneral **A**nesthetic hyper-**S**ensitive response, is located on the X chromosome and encodes a core 49 kDa subunit of mitochondrial ETC complex I that is required for oxidative phosphorylation (Kayser et al. 1999). The *gas-1 (fc21)* mutant allele encodes a single missense mutation, which substitutes a lysine for a highly conserved arginine (Kayser et al. 1999). The *gas-1 (fc21)* C → T mutation was generated through ethyl methanesulfonate (EMS) mutagenesis, which is predicted to have a C:G → T:A transition bias (Kayser et al. 1999; Kim et al. 2006).

Previous work (Hartman et al. 2001; Kayser et al. 2004; Lenaz et al. 2006) and data from the Estes lab show that the *gas-1* mutant exhibits many deleterious phenotypes including: low reproductive fitness, high reactive oxygen species (ROS) levels, altered mitochondrial membrane potential, and low ATP levels, but low and steady rates of pharyngeal pumping (a proxy for chronological age), relative to wildtype.

Complex I is the first and largest component of the mitochondrial ETC (Hunte et al. 2010). It oxidizes NADH—generated from the citric acid cycle and beta-oxidation of fatty acid—and uses the resultant two electrons to reduce ubiquinone to ubiquinol. The redox energy generated from the oxidation of NADH is sufficient to drive four protons across the inner mitochondrial membrane (Wikström 1984). Although there are gaps in our understanding of the exact mechanisms of complex I and the GAS-1 subunit (Wirth et al. 2016), biochemical evidence suggests that GAS-1 funnels electrons through complex I as part of an iron-sulfur (Fe/S) cluster of cofactors (Fassone and Rahman 2012). X-ray crystallography resolution of complex I isolated from bovine heart has further revealed that the 49 kDa GAS-1 subunit creates a binding site cavity in conjunction with the 30 kDa PSST subunit where ubiquinone is ultimately reduced (Vinothkumar et al. 2014; Zickermann et al. 2015; Wirth et al. 2016). The *C. elegans* GAS-1 protein is 83.4% homologous at the amino acid level to human NADH dehydrogenase ubiquinone iron-sulfur protein 2 (NDUFS2) (Kayser et al. 2011). Complex I is also the primary site of ROS production in the ETC (Hirst et al. 2008). Two active sites—NADH and ubiquinone binding sites—in complex I are thought to be primarily responsible for the leakage of electrons onto oxygen, which generate ROS (Genova et al. 2001; Hirst et al. 2008). A mutation in a

highly conserved residue lying adjacent to the active site of ubiquinone binding is therefore likely to have an impact on ROS production—all other things held constant.

Research overview. In the following chapters, I outline my work to determine the capacity of compensatory and generally beneficial mutation to mitigate *gas-1 (fc21)* ETC dysfunction in *C. elegans* nematodes and restore wildtype fitness and phenotypes. In Chapter Two: Phenotypic Recovery Under Unfavorable Population Genetic Conditions, I explore the capacity of new mutations to mitigate *gas-1 (fc21)* ETC dysfunction under MA conditions. The MA process minimizes the power of selection and allows all but the most deleterious mutations to accumulate via genetic drift in the independently evolving MA lines (Halligan and Keightley 2009; Vassilieva and Lynch 1999). To accomplish this, I first performed additional analyses on fitness assay data collected on 48 *gas-1* MA lines evolved in population bottlenecks of one for an average of 43 generations and compared them to their *gas-1* G0 progenitor and a wildtype Bristol N2 control. I confirm that the *gas-1* MA lines experienced partial fitness and phenotypic recovery towards wildtype levels despite extremely unfavorable population genetic condition. The *gas-1* MA lines also exhibited a decrease in life-history trait variance compared to *gas-1* G0, a pattern more consistent with purifying selection rather than genetic drift. I also describe gene ontology (GO) enrichment analyses performed on nuclear genome mutations (SNPs) that were identified from whole-genome sequencing on a subset of five randomly chosen *gas-1* MA lines and N2 MA lines from (Baer 2005; Denver 2012) that suggest a pattern of positive selection in the *gas-1* MA lines. Lastly, I describe work done to functionally characterize fitness and steady-state ROS and ATP levels on a set of four mutations (SNPs)

that arose in one of the *gas-1* MA lines, MA431, and isolated on *gas-1* G0 and wildtype N2 backgrounds. I determined that the mutations were beneficial for fitness on both a *gas-1* G0 and wildtype N2 backgrounds and compensatory for steady-state ROS and ATP levels.

In Chapter Three: Fitness Recovery in Large Populations, I go on to investigate the capacity of new mutations to mitigate *gas-1* (*fc21*) ETC dysfunction under efficient selection conditions. Previous experiments have shown fitness recovery in MA lineages of *C. elegans* subsequently evolved in large population sizes (Estes and Lynch 2003; Denver et al. 2010; Estes, Phillips, et al. 2011); I explored this in recovery in the context of mitochondrial dysfunction. To accomplish this, I first performed a fitness assay on 24 *gas-1* RC lines evolved in large population sizes of roughly 1000 individuals for 60 generations and compared them to their *gas-1* G0 progenitor and a wildtype Bristol N2 control. Interestingly, data from the fitness assay showed no increase in life-history trait means of the *gas-1* RC lines as a group compared to *gas-1* G0; however, a subset of the lines that we observed generating males or being resistant to age-synchronous bleach treatments during experimental evolution experienced robust fitness recovery towards wildtype N2. Male frequency assays performed on a subset of the *gas-1* RC lines including all the lines with observational evidence for males showed that only two of the six *gas-1* RC lines with evidence for males actually had significant levels of males; however, this finding and its caveats are discussed in the chapter. Though, non-disjunction assays performed on the same subset of lines as the male frequency assays revealed no elevated rates of non-disjunction. Finally, I performed gene ontology enrichment analyses on nuclear genome mutations (SNPs) that were identified from whole-genome sequencing

on the *gas-1* RC lines; these analyses were compared with the GO enrichment analyses from the *gas-1* MA lines in Chapter Two: Phenotypic Recovery in the Face of Mutation Accumulation.

Finally, in Chapter Four: Conclusion, I discuss broader conclusions about how the research outlined in this thesis relates to- and broadens the current knowledge base on the phenotypic patterns and mechanisms of ETC disease recovery, and on experimental evolution in general. We also discuss work being done currently to address some of the outstanding questions of our research and possible areas for future work.

CHAPTER 2: PHENOTYPIC RECOVERY UNDER UNFAVORABLE POPULATION GENETIC CONDITIONS

Background

Mutation is the ultimate source of genetic variation in populations. However, most new mutations are deleterious for organismal fitness and can propagate within populations under a broad range of conditions (Crow and Kimura 1970; Eyre-Walker and Keightley 2007; Kondrashov and Kondrashov 2010). When multiplied across the total number of sites within a genome, the cumulative effects of mildly deleterious mutation accumulation can add up to a large impact. If such mutations are not purged by selection, they can lead to a steady decline in fitness over evolutionary time—perhaps posing a special threat to the survival and integrity of small populations (Lande 1994; Lynch, J. Conery, and Burger 1995).

Both endogenous mutagens—such as low fidelity DNA replication—and exogenous mutagens—such as UV radiation—can lead to spontaneous genetic mutations (Baer et al. 2007). Among these potential mutagens are reactive oxygen species (ROS), which can be generated in many enzymatic systems (Holmström and Finkel 2014) and as natural byproducts of cellular respiration in the mitochondrial electron transport chain (ETC) (Murphy 2009). ROS are important cell signaling molecules (Mittler et al. 2011) helping to regulate apoptosis (Fleury et al. 2002), activate MAPK pathways (Jammes et al. 2009), and DNA repair (Vanderauwera et al. 2011). However, they may also be a key promoter of mutation (Hsie et al. 1986; Demple and Harrison 1994; Cooke et al. 2003). ROS-induced mutagenesis has a hallmark signature of DNA damage G:C → A:T transition

mutations and G:C → T:A transversion mutations (Wiseman and Halliwell 1996; Cooke et al. 2003; Cooke and Evans 2007). While the impact of ROS on somatic mutation rates and patterns has been studied (Wiseman and Halliwell 1996; Polyak et al. 1998; Taylor and Turnbull 2005; Garcia et al. 2010; Ameer et al. 2011), there is little evidence for the effect of increased levels of ROS on mutation rates and the distribution of mutational effects across generations because ROS-induced DNA damage studies typically center on somatic cells and repair mechanisms, rather than ROS-induced mutations in the germline (see exception Joyner-Matos et al. 2011; Joyner-Matos et al. 2013; Wernick et al. 2016).

The exact population-level consequences of new mutation (such as mutation meltdown (Lynch et al. 1995b)) depend on the mutation rate and the distribution of new mutations' effect on fitness. The most reliable empirical estimates of both parameters come from mutation-accumulation (MA) experiments wherein replicate lineages of a genotype are bottlenecked in small effective population sizes for many generations, minimizing the power of selection to purge deleterious (reviewed in (Lynch et al. 1999; Halligan and Keightley 2009)). Mutation rates and distributions of effects can be inferred from fitness data on MA lines compared to their ancestor (Keightley 1994; Garcia-Dorado 1997; Shaw 2002; Estes et al. 2003); however, it is now possible to obtain direct estimates of the mutation rate with genome sequencing (Lynch et al. 2008; Haag-Liautard et al. 2008; Denver et al. 2009; Ford et al. 2011; Saxer et al. 2012; Colangeli et al. 2014). With few exceptions (Shaw et al. 2002; Joseph and Hall 2004; Rutter et al. 2010), these experiments have revealed a decrease in per generation mean fitness caused by a net

accumulation of deleterious mutations, with with as much as 10X higher mutation rates in mtDNA as compared to nDNA (Haag-Liautard et al. 2008).

The fact that populations persist at high levels of fitness under this constant pressure of deleterious mutations suggests that other evolutionary processes can ameliorate the phenotypic decline caused by deleterious mutations (e.g., Estes et al. 2011). Three mutational processes able to counteract deleterious mutation accumulation include: 1) reversion mutations that restore wildtype sequences, 2) generally beneficial mutations that improve organismal fitness irrespective of genetic background, and 3) compensatory mutations that specifically mitigate the effects of previously-acquired deleterious mutations through epistasis (Kimura 1985). Unlike generally beneficial mutations, compensatory mutations are beneficial *only* within the context of a particular deleterious mutation and are, by themselves, either neutral or mildly deleterious. The potential for any of these mutation types alters our expectations for the impact of deleterious mutation in populations or (Orr 2000; Poon and Otto 2000a; Lanfear et al. 2014), but since the fitness effects of individual mutations are rarely characterized, the relative importance of beneficial and compensatory epistatic mutations is unknown.

Theoretical and some empirical evidence suggests that mutations have different selective effects depending on molecular genetic context; a mutation with a specific phenotypic effect may be deleterious in one genome and advantageous in another (Wright, 1968). This is in part due to epistatic gene-gene interactions, where the effects of one gene are altered or masked by the presence of a modifier gene (Phillips 2008). An example of this is seen in complex disease traits like Alzheimer's where there are

estimated to be over 100 epistatic interactions where certain genes either promote the development of disease or confer protection (Lobo, 2008). This context-dependence complicates the effect on fitness for newly arising mutations in populations (Poon and Otto 2000). Landscape theory also suggests that the effects of new mutation and the distribution of new mutations' effect on fitness will depend on a population's current fitness. Ascertaining the organismal- and population-level consequences of mutation is therefore difficult in part because the effects of mutation are context-dependent. Experiments with well understood study-systems are necessary to study the effects of new mutation in an evolutionary framework.

The mitochondrial electron transport chain (ETC) offers a compelling framework for studying the accumulation and effects of mutations. Mitochondria are essential eukaryotic organelles that generate the majority of cellular bioenergy (in the form of ATP), but also produce ROS (Murphy 2009). ETC components pump protons across the mitochondrial intermembrane creating a proton motive force, which drives the production of ATP. Because the mitochondrial ETC is highly conserved, vital for energy metabolism in all eukaryotic life, it is an ideal system in which to study the impact of deleterious mutation. Many features of mitochondrial biology—high mutation rate, apparent lack of recombination, uniparental inheritance, and proximity to potentially damaging ROS—could potentially promote faster rates of evolution in mtDNA (Ballard and Whitlock 2004). In addition, since proper ETC functioning relies on the maintenance of favorable mitonuclear epistatic interactions (Blier 2001; Dowling et al. 2008), it is an excellent system for addressing the role of epistasis in maintaining population fitness.

This thesis chapter builds upon previous work by Ph.D. student, Riana Wernick (D. R. Denver Lab at Oregon State University, Corvallis, OR) and M.S. student, Michael Lue (S. Estes Lab at Portland State University). Wernick performed an MA experiment wherein mutations were allowed to accumulate for an average of 43 generations in 48 replicate lineages evolving in population sizes of one using a mitochondrial ETC mutant strain of *C. elegans* nematode, *gas-1* (Wernick et al. 2016). This strain contains a single missense mutation affecting ETC Complex I (Kayser et al. 1999) and exhibits deleterious phenotypes including: low reproductive fitness, high levels of potentially damaging ROS, and low ATP levels alongside slowed pharyngeal pumping rates (a proxy for chronological age) relative to wildtype (Hartman et al. 2001; Kayser et al. 2004; Lenaz et al. 2006; Lue 2015). The main goal of the *gas-1* MA experiment was to assess whether increase levels of endogenous ROS would lead to an increase in genomic mutation rates and how this would affect fitness. Wernick et al. 2016 studied the impact of mitochondrial dysfunction on mtDNA mutation using the MA approach. Lue (2015) explored the impact of MA in the *gas-1* genotype on fitness and mitochondrial function. The MA process was expected to result in further fitness decline due to accumulated deleterious mutations (Kondrashov and Kondrashov 2010), and, perhaps, an increased rate of mutations alongside increasing endogenous steady-state ROS levels compared to wildtype. However, the *gas-1* MA lines experienced partial phenotypic recovery towards wildtype levels (Lue 2015), and whole-genome sequencing and bioinformatic analyses showed the mutation rate was not statistically different than wildtype (Wernick 2016).

Here, I add to this work by analyzing additional fitness assay data on the *gas-1* MA lines and expanding the analyses to report relative fitness, phenotypic and mutational variance components, and per-generation changes in life-history traits. I also report on the phenotypic and molecular characterization of the *gas-1* MA lines nuclear SNPs and compare them to a set of nuclear SNPs that arose in a set of wildtype N2 MA lines (Baer et al. 2005; Denver et al. 2009) using bioinformatic analyses. And finally, I characterize four of the nuclear SNPs isolated from one *gas-1* MA line isolated on both the wildtype N2 and *gas-1* backgrounds to functionally test the nature and selective strength of the mutations. My research allowed us to identify patterns in selection and drift in *gas-1* recovery under MA conditions. This research also advances our understanding of the genetic bases of adaptive evolution under extremely unfavorable conditions and of how mitochondrial dysfunction affects evolutionary dynamics.

Methods

Strains and culture conditions. This thesis utilized MA lines generated from either a *gas-1* mutant or standard laboratory wildtype, Bristol N2, strain of *C. elegans*. The *gas-1* mutant, derived from ethyl methanesulfonate (EMS) mutagenesis (Kayser et al. 1999), was obtained from the Caenorhabditis Genetics Center (University of Minnesota) and repeatedly backcrossed to N2 in an attempt to create an isogenic mutant strain. As detailed in Wernick et al. 2016, a *gas-1* mutant strain, CW152, was backcrossed to N2 for 10 generations during which time PCR and Sanger DNA sequencing were used to track the *gas-1* (*fc21*) allele, a C → T point mutation that replaces a conserved arginine with lysine in the GAS-1 protein. Offspring of a single *gas-1* (*fc21*) hermaphrodite were then used to

initiate 48 MA lines. As previously described (Wernick et al. 2016), lines were maintained by transferring single larval stage 4 (L4) offspring to fresh plates at 4-day intervals (corresponding to one generation for this strain) for an average of 43 generations (range: 35-47 generations), after which time lines were stored cryogenically at -80 °C. Failed bottlenecks occasionally resulted from hermaphrodite sterility or death prior to reproduction; in this event, a sibling from the same generation was chosen from a backup plate to initiate the next generation—a practice that generated variation among lines in the final number of MA generations achieved. An MA line was considered extinct when it could not be reconstituted in this way. This study also utilized data collected on a subset of five *gas-1* MA lines—412, 419, 429, 431, and 438—selected at random for whole-genome sequencing and additional phenotypic analyses. MA lines generated from the wildtype N2 strain were obtained from Charles Baer (University of Florida) who evolved replicate lines of the highly inbred N2 strain in the manner described above for a maximum of 250 generations (Baer et al. 2005). Specifically, this thesis utilized a subset of five N2 MA lines—523, 526, 529, 553, and 574—for which genome sequences (Denver et al. 2009; Denver et al. 2012) and phenotypic data (Denver et al. 2009; Clark et al. 2012; Joyner-Matos et al. 2013; and Andrew et al. 2015) were available.

Toward the goal of isolating and individually characterizing a set of candidate beneficial or compensatory SNPs that arose on a single chromosome (III) within one particular MA line, this line (MA431) was backcrossed to N2 for 7 generations by Dana Howe (D. R. Denver Lab at Oregon State University, Corvallis, OR) following the approach used in Wernick et al. 2016. At each generation, DNA was extracted from the

parental MA431 hermaphrodite (n = 30-44 nematodes) for PCR and Sanger DNA sequencing at three SNP locations within the *rheb-1*, *daf-2*, and *sel-2* genes on chromosome III. Only offspring from hermaphrodites that retained remained heterozygous at least one of these SNPs were used in the next round of crossing. Although our primary aim was to isolate each chromosome III SNP individually, we were unsuccessful at maintaining strains that did not contain all three (see Results for MA431 SNP linkage). At any generation where we found a strain that lost one of the three heterozygous chromosome III SNPs, the other two heterozygous SNPs on chromosome III would be lost at the next backcross generation or during a round of selfing in an attempt to make at least one of the remaining two SNPs homozygous. After seven rounds of backcrossing, the hermaphrodites containing MA431 mutations were allowed to self and tested to find homozygotes, resulting in a strain referred to hereafter as N2^{MA431}. Finally, the MA431 chromosome III homozygous mutant lines were crossed with our near-isogenic *gas-1* mutant strain and offspring allowed to self to produce a strain that contained only the chromosome III SNP mutations on the *gas-1* background, *gas-1*^{MA431}.

Prior to all experiments, strains were allowed to recover from freezing for two to three generations at 20 °C on Nematode Growth Medium-Lite (NGML) plates containing 20 µg/mL streptomycin and seeded with *Escherichia coli* OP50-1. For ATP and ROS assays, lines were age synchronized by standard bleach treatment prior to each assay.

Life history assays. I assayed daily offspring production and lifespan for an additional block of the *gas-1* MA lines (Lue 2015), and the N2^{MA431} and *gas-1*^{MA431} lines following established methods (e.g., Estes et al. 2003; Estes et al. 2011). Owing to difficulty in

working with the poorly-performing *gas-1* MA lines, life-history assays were previously conducted in four experimental blocks in which different subsets of *gas-1* MA lines were assessed alongside N2 and *gas-1* ancestral (hereafter “*gas-1* G0” for generation 0) controls (Lue 2015). A fifth and sixth block were conducted to increase the sample size of the five *gas-1* MA lines selected for whole genome sequencing; I updated the original data in Lue 2015 with these assay data. Prior to each assay block, lines were segregated into two replicates during recovery from freezing and one adult hermaphrodite (“parent” below) from each replicate was transferred to an individual plate to lay eggs for five hours, creating an age-synchronous cohort. A roughly equal number of offspring from each hermaphrodite, with slight variation due to number of eggs laid, were then transferred to individual seeded plates, resulting in a total of 9-10 individuals from each *gas-1* MA line (or 9-26 from the subset that underwent DNA sequencing) and 10-25 control individuals assayed in each block. A second fitness assay was performed on the N2^{MA431} and *gas-1*^{MA431} strains alongside N2 and *gas-1* G0 controls; this assay included 40 replicates each of the N2^{MA431}, *gas-1*^{MA431}, and ancestral control strains.

Individuals were transferred daily until cessation of egg laying to separate parents from offspring. Plates were incubated at 20 °C for 36 hours to allow eggs to hatch; resulting offspring were killed with one drop of 0.5 M sodium azide and counted directly by contrast staining plates with toluidine blue. We also recorded the number of unhatched and presumably unviable eggs, which are easily distinguished from unfertilized eggs under a light microscope. Lifespan was calculated as days lived from the egg stage; survival rates were calculated from the same dataset. Across both assays, 19 worms died

due to transfer accidents or desiccation after crawling up the side of Petri plates and were omitted from analysis.

The above data were used to calculate mean relative fitness (ω) of the *gas-1* mutant compared to N2, and of each *gas-1* MA line compared to the *gas-1* G0 ancestor. Relative fitness of each individual was computed following (Peters et al. 2003) as: $\omega = \Sigma e^{-rx} l(x) m(x)$, where $l(x)$ is the number of worms surviving to day x , $m(x)$ is the fecundity at day x , and r is the mean intrinsic population growth rate of the assay-specific N2 or *gas-1* G0 control as appropriate. The latter was calculated by solving Euler's equation for r from the equation $\omega = \Sigma e^{-rx} l(x) m(x) = 1$ using an average value of $l(x) m(x)$ for each block-specific control. I set $x = 4.75$ on the first reproductive day (cf. Vassilieva et al. 2000).

Life-history data analysis. I estimated the per generation change in the mean, ΔM , of each trait as the generalized linear regression coefficient of trait value scaled as a fraction of the block-specific *gas-1* G0 mean on line-specific MA generation number. For the MA lines, within- and among-line components of variance were calculated using restricted maximum likelihood (REML) with variance estimates constrained to be non-negative (except for when calculating confidence intervals). I tested the model $trait\ mean = \mu + block + MA\ line(block) + parent[MA\ line(block)] + \epsilon$, wherein all effects are random—and $MA\ line(block)$ represents among-line (genetic) variance and ϵ represents the within-line (environmental) component of variance, V_E . From the same model, I created a frequency distribution of *gas-1* MA line aggregate mutational effects on fitness relative to *gas-1* G0 with least squares means for the *gas-1* MA lines derived from the $MA\ line(block)$ term

using the expected mean squares (EMS) method as applied in JMP12 (SAS). For N2 and *gas-1* G0 lines, we tested the model $trait\ mean = \mu + block + parent(block) + \varepsilon$, wherein all effects are random and *parent(block)* represents among-line (or pseudoline) variance. The per-generation change in genetic variance owing to mutation accumulation, the mutational variance V_M , was calculated for each trait as the difference in the among-line components of variance of the *gas-1* G0 control and the MA lines divided by $2t$, where t = the maximal number of MA generations. The per generation change in among-line variance, V_b , was calculated for the *gas-1* MA lines by dividing the among line variance for each line by the average number of generations, 42.8. I used the Cox proportional hazards method to compare age-specific mortality of the N2^{MA431} and *gas-1*^{MA431} strains and their N2 and *gas-1* G0 controls, with *strain* and *parent(strain)* as main effects. Selection coefficients of the MA431 SNPs on each genetic background were calculated as the difference in relative fitness between N2^{MA431} and N2 and between *gas-1*^{MA431} and *gas-1* G0.

Steady-state ROS levels. I quantified *in vivo* steady-state reactive oxygen species (ROS) levels of the N2^{MA431} and *gas-1*^{MA431} lines compared to N2 and *gas-1* G0 strains using established fluorescence microscopy methods (Hicks et al. 2012; Joyner-Matos et al. 2013; Smith et al. 2014; Lue 2015). ROS levels of the five sequenced *gas-1* MA lines and N2 MA lines were previously reported (Lue 2015 and Joyner-Matos et al. 2013, respectively). Briefly, age-synchronized worms were incubated for 24 hours in 1 mL of 10 μ M MitoSOX Red mitochondrial superoxide indicator (Molecular Probes, Inc., Eugene, OR) in M9. Strains were washed and transferred to plates with unlabeled *E. coli* for one

hour before imaging. A second set of unlabeled strains served as control. Worms were paralyzed with 5 M levamisole immediately prior to imaging. I collected 5 μ m z-stack images of the terminal pharynx using a wide-field Zeiss Axio Imager 2 mounted with a monochrome Axiocam CCD camera at OHSU. A different camera and microscope setup was used than Leu 2015 owing to a core facility equipment change. Three imaging blocks were completed for the N2^{MA431} and *gas-1*^{MA431} lines with 10–20 labeled worms and roughly 5 control worms from each line imaged during each session. Terminal bulbs in each z-stack image were manually encircled and average maximum fluorescence intensity measured using ImageJ software (NIH, Bethesda, MD). Final ROS levels were calculated as the difference between average maximum pixel intensities (across z-stacks for each individual) for labeled and line-specific control worms.

Steady-state ATP levels. I quantified ATP content for 4-5 independent samples from each of the N2^{MA431} and *gas-1*^{MA431} lines compared to N2 and *gas-1* G0 strains following methods adapted from Van Raamsdonk et al. 2010 and Yang and Hekimi 2010. ATP data for the *gas-1* MA and N2 MA lines (Leu 2015). For each sample, roughly 200 age-synchronized young adult worms were washed and worm cuticles were broken via successive freeze-thaw cycles. Worms were then boiled for 15 minutes to arrest ATPase activity. Each solution was sonified (20% amplitude, twice for 12 seconds) and centrifuged at 10,000 x g for 12 minutes. I used 90 μ L of supernatant to quantify standing ATP levels for each strain with a luminescence ATP detection kit (Invitrogen, Carlsbad, CA) and a TECAN Infinite M200 Pro plate reader (Tecan Group, Männedorf, Switzerland). ATP

measurements were normalized by protein content measured with a Pierce BCA Protein Assay Kit (Thermo Fisher Scientific, Waltham, MA).

Identification and characterization of *gas-1* RC line mutations; mutation rate analysis.

The Dee Denver Lab at Oregon State University (Corvallis, OR) whole-genome sequenced (Illumina MiSeq) twelve nematode strains including five randomly chosen *gas-1* MA lines, five previously-studied N2 MA lines (Baer et al., 2005), and their respective ancestors, N2 and the backcrossed *gas-1* G0. Samples were prepared from first larval (L1) stage nematodes as described in (Wernick et al. 2016). Candidate SNPs in *gas-1* MA lines were identified as variants from the *C. elegans* reference genome (WS242) and our wild-type N2 lab strain. The nuclear mutation rate was calculated from pooled *gas-1* MA lines using the equation $[\mu_{bs} = m / (LnT)]^{1/2}$ where μ_{bs} is the base substitution rate (per nucleotide site per generation), L is the number of MA lines, m is the number of observed mutations, n is the number of nucleotide sites, and T is the time in generations as previously described (Denver et al., 2009). The resulting data were also used to characterize the *gas-1* genetic background compared to the N2 strain, and to investigate the possibility of a translocation mutation within MA431 (R. Wernick, D. Howe, and D. Denver, unpubl.).

Gene ontology enrichments. I examined gene ontology (GO) enrichments for the genes in the *gas-1* interactome and all nuclear DNA SNPS that arose within protein coding regions of *gas-1* MA and the Baer N2 MA lines within three functional domains—biological process, cellular component, and molecular function. I used GoMiner (Zeeberg et al. 2003, application build 457) and the GO MySQL Database (MySQL 3.7.13, Oracle Corporation, Cupertino, CA; GO database build 2016-06-07, geneontology.org) to

calculate GO enrichments. GoMiner reports a two-sided Fisher's exact p value based upon the number of genes in a category and a false discovery rate (FDR)-corrected p value for multiple testing (Zeeberg et al. 2005); however, the FDR can be overly conservative when the sample size or number of discoveries is small; i.e., less than 100 (Tong and Zhao 2008; Zeeberg et al. 2005), so I based my interpretations on results of the Fisher's exact tests with $\alpha = 0.05$. I then determined which broader-level functional "GO slim" categories, as defined by the GO Consortium (Version 1.2, 2012-09-21, geneontology.org), were enriched using CateGORizer (Hu et al. 2008). CateGORizer classifies enriched GO terms into their respective GO slim categories, giving a coarser-scale overview of enrichment patterns. GO slim categories were then checked against the GoMiner output to assess statistical significance.

Interactome analysis of gas-1 and N2 MA line mutations. In coordination with R. Wernick, I determined if *gas-1* MA line mutations interacted within 2-degrees of the *gas-1* gene as compared to mutations that arose on a wildtype N2 background. GeneOrienteer version 2.25 (Zhong and Sternberg, 2006) was used to construct interactomes, the complete list of all genes predicted to directly interact with a gene of interest. GeneOrienteer computationally integrates interactome (known gene interactions), gene expression, phenotypic, and functional annotation data from *C. elegans*, *Drosophila melanogaster*, and *Saccharomyces cerevisiae* to predict genome-wide interactions in *C. elegans* (Zhong and Sternberg, 2006). Due to intrinsic error present in data from disparate organisms, a likelihood ratio test is employed in GeneOrienteer to assess the likelihood of two genes interacting in *C. elegans*. A filter is applied to minimize false positives and

exclude genetic interactions among genes in the same orthologous group. A likelihood ratio assigns a weighted score to five features: (1) identical anatomical expression, (2), phenotype, (3) functional annotation, (4) microarray coexpression, and (5) the presence of interlogs. Gene interactions below a minimum predicted likelihood ratio score of 4.6 ($= 2\ln(10)$) were excluded from analysis due to low significance when converted to X^2 test statistics. At the time of analysis, GeneOrienteer showed 500,303 *C. elegans* gene-gene interactions to have interactions values at or above 4.6 (Zhong and Sternberg, 2006). We queried the *C. elegans* interactome to determine the interactions containing *gas-1*. The query returned 181 genes predicted to interact with *gas-1* directly within one-degree. We then queried the *C. elegans* interactome with the 181 genes predicted to directly interact with *gas-1* to determine the genes predicted to interact within two-degrees of *gas-1*. We identified 4,008 unique genes that are predicted to interact with the 181 direct *gas-1* interactors. The comprehensive list of first- and second-degree *gas-1* interacting genes, totaling 4,189, is referred to hereafter as the “*gas-1* centric interactome”. We identified all genic mutations within the *gas-1* and N2 MA line sets that were located in genes predicted to reside within the *gas-1*-centric interactome. R. Wernick then ran Monte Carlo Maximum Likelihood simulations to assess the likelihood that the number of *gas-1* MA and Baer N2 MA genic mutations would overlap with *gas-1* centric interactome genes by chance alone (Wernick 2016).

Results

Evidence for selection during *gas-1* MA. Though not previously reported, eight of the original 48 *gas-1* MA lines (16.7%) went extinct prior to the 50 generation cutoff, as

compared to only one extinction among 48 MA lines initiated from N2 (2%) during 125 generations of evolution (Joyner-Matos et al. 2011). And the frequency of failed bottlenecks occurring during the *gas-1* MA experiment was 14.2%—twice that reported for the Baer N2 MA lines (7.1%; Joyner-Matos et al. 2011). These patterns are indicative of stronger purifying selection operating at some level (gametic, zygotic, individual) in *gas-1* MA lines. Additionally, *gas-1* MA lines exhibited an increased number of fertilized but ultimately unhatched eggs compared to ancestral controls. The difference in unhatched egg counts among N2, *gas-1* G0 and MA groups ($F_{(2, 677)} = 8.9345$, $p < 0.0001$) was due to the *gas-1* MA line group having a small but significantly greater number of unhatched eggs (3.68 ± 0.31) than either *gas-1* G0 (1.51 ± 0.41) or N2 (1.09 ± 0.18) (Tukey HSD; $\alpha = 0.05$, $p < 0.01$). This pattern suggests that postzygotic selection was occurring in the *gas-1* MA lines.

***gas-1* G0 characterization and evolution of life-history trait means.** There were no significant effects of assay block on any life-history trait (not shown) with the addition of block six to the previously conducted fitness assays (Leu 2015). As expected based on previous studies of the CGC *gas-1* strain (Morgan and Sedensky 1994; Kayser et al. 1999), and the previous fitness assays performed in our lab (Leu 2015), our backcrossed *gas-1* G0 strain exhibited extremely low offspring production compared to wildtype N2 (Figure 2.1; Table 2.1, Table 2.2), resulting in significantly lower relative fitness ($t_{(148)} = 9.954$, $p < 0.0001$; Table 2.1). Furthermore, and concordant with Leu 2015, *gas-1* G0 reproductive maturity and peak reproductive output were delayed by a full day compared to N2 (Figure 2.1). In contrast to previous reports of *gas-1* being short-lived (Kayser et al.

2004), our *gas-1* G0 lived longer than N2 (Table 2.1; Tukey's HSD, $p < 0.01$); average lifespan for our *gas-1* G0 strain (16.3 days) was >6 days longer than that reported for the CGC strain (9.8 days) (Kayser et al. 2004). Major caveats to this result, however, include that our N2 performed worse than expected for this assay, living an average of ~13 days (Table 2.1). Although this value falls within the normal range reported for N2, 12–18 days (Riddle et al. 1997), variation among particular N2 lab strains (and across assays) is common, and our lab N2 has previously lived longer at 16.8 days (Smith et al. 2014) and 17.2 days (Estes et al. 2011). Secondly, performance of our *gas-1* G0 strain differed substantially between this and the second assay the reported below. Neither difference in the frequency of “bagging” (where internal hatching of eggs kill the parent) nor of unhatched eggs can explain the reduced lifespan of N2 for this assay. Instances of bagging in N2 (12.38%) were no more common than we have previously observed for this strain (not shown), but slightly more common than observed for *gas-1* G0 (10.23%; NS). Similarly, we found no evidence for an increased presence of unhatched eggs in N2 for this assay compared to previous assays (not shown), and N2 did not differ from *gas-1* G0 in this regard (reported below).

Turning to the performance of the *gas-1* MA lines, relative fitness and total reproduction were significantly higher on average in the MA lines compared to their *gas-1* G0 ancestor (Figure 2.2; Table 2.1, Table 2.2), leading to positive values for ΔM (Table 2.1). Least squares means for relative fitness increased ($\omega > 1$) for 18 *gas-1* MA lines and decreased ($\omega < 1$) for 30 *gas-1* MA lines compared to *gas-1* G0 ($\omega = 1$) (Figure 2.2). These differences were due to MA lines producing, on average, significantly more offspring on

days 1, 3, 4, and 5 of their reproductive period (Figure 2.1; Table 2.2). On the other hand, MA lines lived 1.68 days less on average compared to *gas-1* G0, but this difference was not significant and ΔM for lifespan is not distinguishable from zero (Table 2.1). The prevalence of bagging in the MA lines (8.2%) was no different from that of N2 or *gas-1* G0 (χ^2 test), and thus did not significantly impact MA line lifespan.

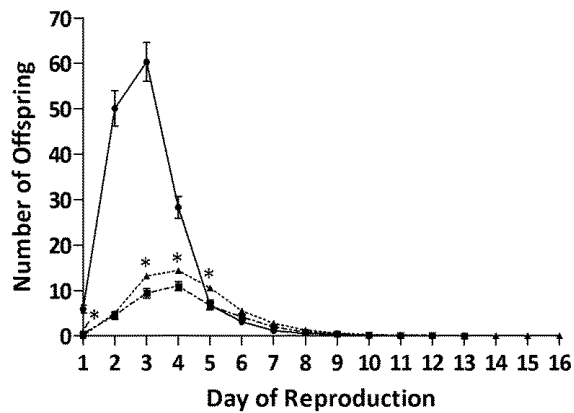


Figure 2.1. Reproductive schedules. Daily offspring production of N2 (solid line), *gas-1* G0 (dashed line, squares), and the combined *gas-1* MA lines (dashed line, triangles). Error bars are 1 S.E.M. Summary statistics are reported in Table 2.2.

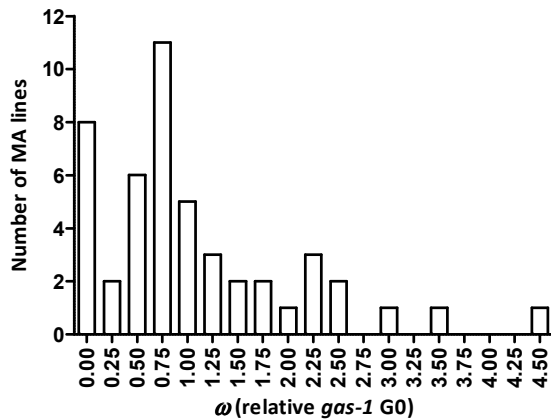


Figure 2.2. Frequency histogram of *gas-1* MA relative fitness. Histogram of the 40 *gas-1* MA lines' fitness (w) relative to *gas-1* G0 (equal to 1.0) using least squares means including the eight *gas-1* MA lines that went extinct.

Evolution of life-history trait variances. Phenotypic variance among ancestral N2 or *gas-1* G0 control lines did not differ from zero for any trait, although the estimates for *gas-1* G0 are higher for relative fitness and total productivity, and lower for lifespan, compared to those for N2 (Table 2.1). This pattern is mirrored in the estimates of within-line (environmental) variance (Table 2.3). Significant per-generation changes in the among-line (genetic) variance, V_b , were detected for productivity and relative fitness of *gas-1* MA lines (Table 2.1). However, point estimates for V_b were uniformly lower for *gas-1* MA lines compared to those for the *gas-1* G0 ancestor for all traits, and V_b for lifespan was indistinguishable from zero (Table 2.1). If we had corrected for the (non-significant) among-(pseudo)line variance observed for the *gas-1* G0 control as some previous MA studies have done (e.g., Joyner Matos et al. 2011), we would have detected no significant change in genetic variance for the MA lines. Accordingly, mutational variance, V_M , estimates are low for all traits, and slightly negative for lifespan (Table 2.1). Finally, a roughly similar pattern of reduced variance was observed for V_E , the per-generation within-line variance, meaning that average MA line performance became more uniform compared to that of its ancestor.

Table 2.1. Summary statistics for life-history traits.

	<i>W</i>	<i>ω</i>	<i>Lifespan</i>
N2	151.6 ± 10.03	1.000 ± 0.053	12.97 ± 0.429
<i>gas-1</i> G0	38.40 ± 3.277	0.225 ± 0.057	16.33 ± 0.812
<i>gas-1</i> G0 [‡]	38.40 ± 3.277	1.000 ± 0.085	16.33 ± 0.812
<i>gas-1</i> MA [‡]	50.58 ± 1.647	1.260 ± 0.058	14.65 ± 6.146
ΔM (x 10 ⁻³)	127.9 ± 38.20	5.757 ± 1.317	-0.211 ± 0.865
<i>V_bgas-1</i> G0	76.74 ± 84.81	0.054 ± 0.059	7.579 ± 7.346
<i>V_bgas-1</i> MA	5.191 ± 1.824	0.010 ± 0.004	0
<i>V_Egas-1</i> G0	628.6 ± 105.7	0.014 ± 0.002	50.26 ± 8.542
<i>V_Egas-1</i> MA	20.58 ± 1.439	0.017 ± 0.001	0.838 ± 0.059
<i>V_M</i>	-1.547	0.004	-0.081

All values are means ± 1 S.E.M; [‡] denotes fitness relative to *gas-1* G0; fitness is otherwise reported relative to N2.

Table 2.2. Daily offspring production of N2, *gas-1* G0, and *gas-1* MA lines.

Reproductive Day	N2	<i>gas-1</i> G0	<i>gas-1</i> MA	<i>p</i> value
1	(80) 5.925 ± 0.883	(88) 0.511 ± 0.227	(512) 0.055 ± 0.022	< 0.0001
2	(79) 50.13 ± 3.888	(88) 4.500 ± 0.792	(512) 5.010 ± 0.313	0.5364
3	(78) 60.26 ± 4.269	(88) 9.409 ± 1.016	(511) 13.19 ± 0.527	0.0049
4	(76) 28.33 ± 2.408	(87) 11.03 ± 0.994	(497) 14.40 ± 0.526	0.0113
5	(74) 6.811 ± 1.046	(83) 6.675 ± 0.804	(471) 10.62 ± 0.491	0.0013
6	(69) 3.130 ± 0.563	(81) 4.418 ± 0.697	(437) 5.501 ± 0.324	0.0959
7	(57) 1.211 ± 0.342	(74) 1.986 ± 0.422	(413) 2.782 ± 0.200	0.1154
8	(52) 0.577 ± 0.170	(72) 0.833 ± 0.234	(383) 1.269 ± 0.127	0.1620
9	(43) 0.395 ± 0.153	(66) 0.439 ± 0.140	(354) 0.636 ± 0.085	0.3418
10	(35) 0.029 ± 0.029	(58) 0.224 ± 0.110	(315) 0.260 ± 0.062	0.8122
11	(32) 0.000 ± 0.000	(51) 0.118 ± 0.067	(287) 0.132 ± 0.040	0.8806
12	(29) 0.034 ± 0.034	(45) 0.067 ± 0.067	(255) 0.047 ± 0.016	0.6830
13	(21) 0.000 ± 0.000	(41) 0.049 ± 0.034	(210) 0.038 ± 0.018	0.8028
14	(18) 0.000 ± 0.000	(39) 0.000 ± 0.000	(192) 0.010 ± 0.007	0.5241
15	(13) 0.000 ± 0.000	(34) 0.000 ± 0.000	(163) 0.012 ± 0.009	0.5187
16	(9) 0.000 ± 0.000	(33) 0.000 ± 0.000	(136) 0.007 ± 0.007	0.6237
Total	(80) 151.6 ± 10.03	(88) 38.40 ± 3.277		< 0.0001
		(88) 38.40 ± 3.277	(512) 50.58 ± 1.647	0.0055
			(512) 50.58 ± 1.647	< 0.0001

All values in parentheses are sample sizes; others are means ± 1 S.E.M. The last column shows *p* values associated with two-sided *t*-tests comparing *gas-1* G0 and average *gas-1* MA production. Wilcoxon Rank sums tests were used for pairwise comparisons of total reproduction (bottom 3 rows).

Table 2.3. Variance components for N2 and *gas-1* G0 strains.

Trait	Among-line variances		Within-line variances	
	N2	<i>gas-1</i> G0	N2	<i>gas-1</i> G0
<i>W</i>	2070 (-4084–8225)	76.74 (-89.49–243.0)	4854 (3590–6931)	628.6 (463.7–900.6)
<i>ω</i>	0.034 (-0.041–0.081)	0.007 (-0.011–0.026)	0.270 (0.201–0.380)	0.061 (0.044–0.092)
<i>Lifespan</i>	0.864 (-1.959–3.687)	7.579 (-6.819–21.98)	10.85 (8.265–14.87)	50.26 (36.97–72.31)

Values are among- and within-line variance estimates for life-history traits with lower and upper 95% confidence intervals in parentheses.

Identification and characterization of *gas-1* MA line nuclear mutations; mutation rate

analysis. An overview of whole-genome sequencing and bioinformatics methods used to identify nuclear SNPs in the *gas-1* MA lines was previously detailed (Wernick 2016; and see brief overview in Methods). A total of 40 mutations were detected; of these, 10 occurred in introns, 13 in exons, and the remaining 17 mutations were intergenic (Wernick 2016). Of the 13 exonic mutations, 4 were synonymous and 9 were nonsynonymous. There was no significant difference between the proportion of synonymous to nonsynonymous mutations between the *gas-1* MA and N2 MA values ($p = 0.92$, X^2 test). The mutation rate calculated from pooled data from the five whole-genome sequenced *gas-1* MA lines was $2.13 (\pm 0.34) \times 10^{-9}$ and was statistically indistinguishable (within two standard errors) from the rate previously reported for N2, $2.7 (\pm 0.40) \times 10^{-9}$ (Denver et al., 2009; Wernick 2016).

GO term enrichment. Significant differences between GO terms enriched for *gas-1* MA line and Baer N2 MA line mutations were revealed by our analyses (Appendices A and B, File S2; Figure 2.3). Indeed, no significantly enriched GO slims (see Methods) were shared between the *gas-1* MA and Baer N2 MA line groups (Appendices A and B, Figure 2.3),

although the total number of enrichments was larger for the N2 lines. The latter finding is unsurprising given the greater number of MA generations experienced by the N2 lines. For the *gas-1* MA lines, two GO slim categories were significantly enriched: the biological process of transport (GO:0006810), and the cellular component of plasma membrane (GO:0005886); no molecular function categories were significantly enriched (Figure 2.3). Four *gas-1* MA line genic mutations were classified under plasma membrane; those in: *smf-3*, *daf-2*, *twk-3*, and *rheb-1* genes. These four mutations were also classified under transport. In contrast, twelve GO slims were significantly enriched in the N2 MA lines: biological processes for cell organization and biogenesis (GO:0071840), cell cycle (GO:0007049), and embryonic development (GO:0009790); cellular components for cell (GO:0005623), intracellular (GO:0005622), cytoplasm (GO:0005737), cytoskeleton (GO:0005856), microtubule organizing center (GO:0005815), vacuole (GO:0005773), and lysosome (GO:0005773); and molecular functions for catalytic activity (GO:0003824) and motor activity (GO:0003774) (Figure 2.3). In addition to the lack of overlap between enriched *gas-1* and N2 MA line GO slims (Figure 2.3), we found that all 12 GO slim categories enriched among N2 MA lines overlapped those enriched in the entire *gas-1* interactome set (containing 80 significantly enriched GO slim categories (Appendix C)) while only transport (GO:0006810) was shared between the *gas-1* MA lines and the *gas-1* interactome (Appendices A, B, and C). A complete accounting of GO enrichment can be found in Appendices A, B, and C.

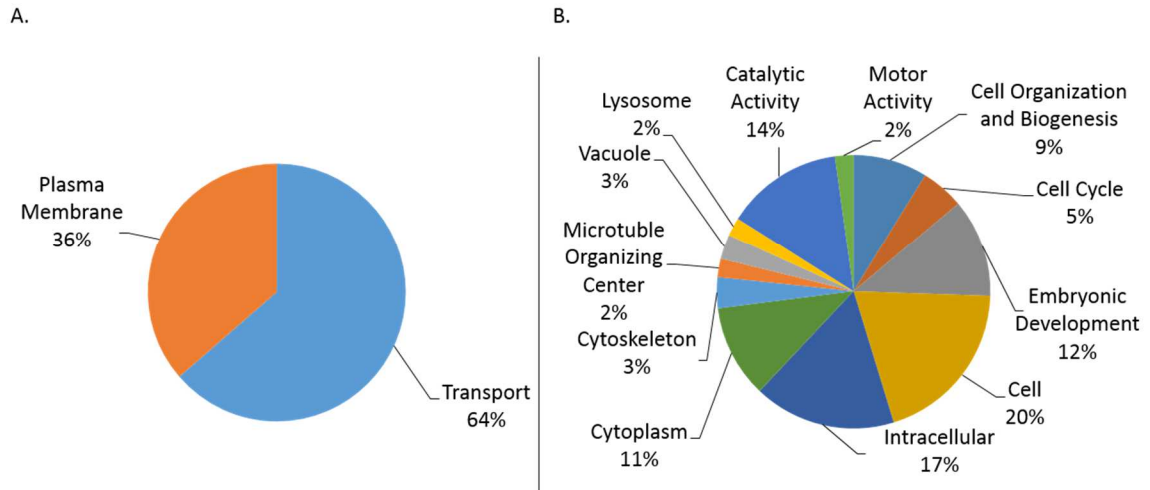


Figure 2.3. Proportions of GO slim categories significantly enriched among the five sequenced *gas-1* MA lines (A) and five N2 MA lines (B). Non-overlapping sets of GO slims were in enriched in *gas-1* and N2 lines, which experienced maxima of 47 and 250 generations of mutation accumulation, respectively. See text for details on analysis and interpretation.

Analysis of *gas-1* interactions. Interactome analyses provided a different view into the patterns and possible functional consequences of mutations fixed within *gas-1* MA lines as compared to those fixed in the Baer N2 MA line subset. R. Wernick evaluated the extent to which mutations acquired by each set of MA lines interacted with genes in the *gas-1* interactome using GeneOrienteer (see Methods). Of the 23 genic mutations found within sequenced *gas-1* MA lines, eight had interactions within two-degrees of the *gas-1* gene: *ZC8.6*, *rheb-1*, *daf-2*, *sel-2*, *smf-3*, *rpi-2*, *W02B8.2* and *alh-2*. One of these, *alh-2*—encoding a mitochondrial aldehyde dehydrogenase—directly interacts with the *gas-1* gene. Of the 60 genic mutations found within sequenced Baer N2 MA lines, 12 were located in genes predicted to interact with the *gas-1*-centric interactome. The ratio of *gas-1* MA mutations in *gas-1* centric interactome genes to the total number of *gas-1* MA genic mutations was significant ($p = 0.005$, χ^2 test) (Wernick 2016); this was not true for the Baer N2 MA line genic mutations.

MA431 SNP linkage. The above bioinformatic analyses revealed several notable SNPs to be contained within one *gas-1* MA line, MA431. In coordination with D. Howe, I sought to isolate and individually characterize these mutations on both *gas-1* G0 and wildtype N2 backgrounds. In particular, among the nine total SNPs discovered within MA431, three affected genes within the *gas-1*-centric interactome (above): *rheb-1*, *daf-2*, and *sel-2*. All reside on chromosome III separated by approximately: 8.5 cM (*rheb-1* and *daf-2*), 3.5 cM (*rheb-1* and *sel-2*), and 11.0 cM (*daf-2* and *sel-2*) (Wormbase, version WS254), meaning we could expect approximately 8.5%, 3.5% and 11% of chromatids from each cross to be recombinant for each pair, respectively. Contrary to this expectation, we observed no recombinant genotypes over the course of the seven-generation backcrossing experiment (see Methods). While this result could be indicative of beneficial epistasis between or among these particular SNPS, our power to detect recombinant genotypes was limited by the small number of offspring ($n = 30-44$) we could practicably sample and genotype at each backcrossing generation. Thus, we cannot rule out scenarios of hitchhiking or linkage among neutral mutations. We were, however, successful in separating these three SNPs from all others except one; a fourth SNP on chromosome V remained linked to the chromosome III set even after seven rounds of backcrossing. This SNP resided within *C04E12.10* (Wormbase, version WS254), an ortholog of human NGLY1 (N-glycanase 1) (Camon, 2003; Mulder, 2003) resulting in replacement of an asparagine with a histidine. We found no evidence that this observation resulted from a translocation event occurring within *gas-1* G0 or MA431 (see Methods); this observation is therefore consistent with the existence of a beneficial- or reciprocal sign epistatic interaction(s)

between the chromosome V SNP and some number of the chromosome III SNPs—or a mtDNA mutation or other nDNA mutation type not considered here. Following our failure to isolate individual SNPs for characterization, I elected to assess the phenotypic effects of the MA431 four-SNP set on the *gas-1* G0 background (strain *gas-1*^{MA431}) and the N2 background (strain N2^{MA431}).

MA431 SNP characterization: Life histories. The four MA431 SNPs increased relative fitness on both *gas-1* G0 and wildtype N2 backgrounds (Figure 2.4), exhibiting a slightly greater selective benefit ($s = 0.065$) on the N2 background (Figure 2.4A) than on the *gas-1* G0 background ($s = 0.029$) (Figure 2.4B). Fitness of *gas-1* G0 and *gas-1*^{MA431} pair of strains was dramatically reduced compared to the N2 and N2^{MA431} pair (Tukey HSD, $\alpha = 0.05$). This between-pair difference overshadowed the small but significant differences within each pair. Wilcoxon rank sums tests showed that N2^{MA431} had improved fitness compared to N2 ($X^2_{(1)} = 5.16, p < 0.02$) (Figure 2.4A) and that *gas-1*^{MA431} had improved fitness than *gas-1* G0 ($X^2_{(1)} = 7.34, p < 0.01$) (Figure 2.4B).

The improved fitness of MA431-SNP containing strains was partly owed to an increase in early-life reproduction compared to their *gas-1* G0 or N2 controls (Figure 2.5; **Table 2.4**). Specifically, *gas-1*^{MA431} produced significantly more offspring on the first two days of its reproductive period compared to *gas-1* G0, while N2^{MA431} produced significantly more than N2 on the first day of reproduction (Table 2.4). Conversely, there was no difference between either pair with respect to total reproduction (Table 2.4); note that N2 “catches up” on the second reproductive day, producing significantly more

offspring than $N2^{MA431}$ (Figure 2.5). Strains also varied with respect to age-specific survival (Figure 2.6) and mortality risk (log-rank $\chi^2_{(3)} = 19.29$, $p < 0.0001$) such that $gas-1^{MA431}$ mortality was reduced compared to $gas-1$ G0 (risk ratio: 0.546, 95% C.I. = 0.347–0.856, $p < 0.008$), but $N2^{MA431}$ mortality was indistinguishable from that of N2. No differences in average lifespan (mean \pm S.E.M.) were detected among $gas-1^{MA431}$ (14.67 ± 0.773), $gas-1$ G0 (12.29 ± 0.613), $N2^{MA431}$ (15.80 ± 0.673), or N2 (17.53 ± 0.586).

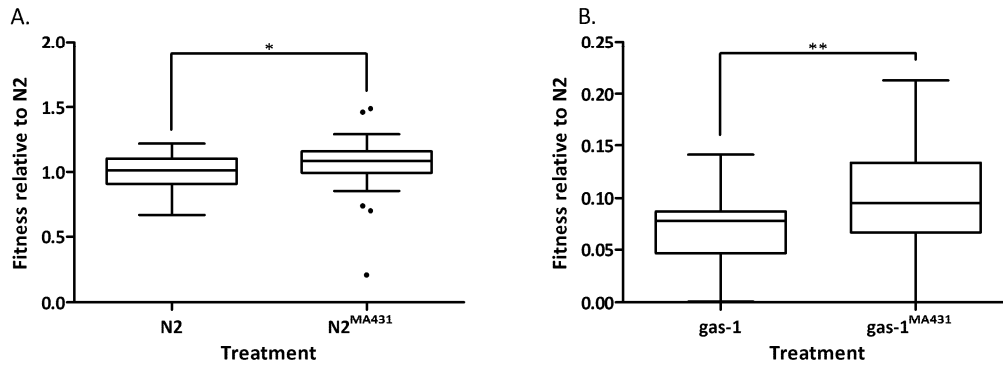


Figure 2.4. Relative fitnesses for MA431 SNP-containing lines and ancestral controls. Fitness of MA431 backcrossed lines relative to N2 reported (mean \pm 1 S.E.M.); note the different y-axis scales in each panel: (A) wildtype N2 (1.000 ± 0.021) and $N2^{MA431}$ (1.065 ± 0.033); (B) $gas-1$ G0 (0.068 ± 0.005) and $gas-1^{MA431}$ (0.097 ± 0.008). * and ** denote statistical significance at $p < 0.01$ and < 0.05 levels, respectively (see text).

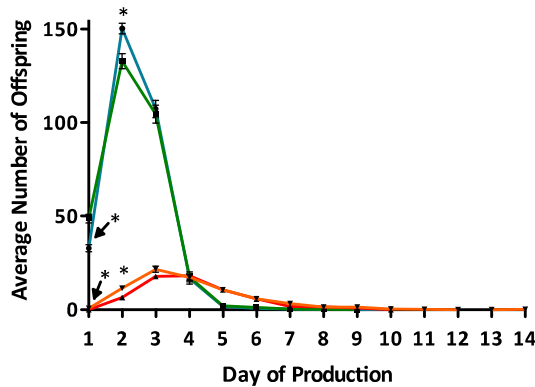


Figure 2.5. Reproductive schedule for MA431 SNP-containing lines and ancestral controls. Average daily offspring production for N2 (blue), $N2^{MA431}$ (green), $gas-1$ G0 (red), $gas-1^{MA431}$ (orange). Error bars = 1 S.E.M. Summary statistics reported in (Table S5).

Table 2.4. Comparisons of daily progeny production of N2, N2^{MA431}, gas-1 G0, and gas-1^{MA431}.

Reproductive Day	N2	N2 ^{MA431}	p value	gas-1 G0	gas-1 ^{MA431}	p value
1	(40) 32.85 ± 1.888	(40) 48.70 ± 2.260	< 0.0001	(41) 0.000 ± 0.000	(39) 0.718 ± 0.244	0.0034
2	(40) 150.4 ± 2.843	(40) 133.0 ± 4.072	0.0008	(41) 6.488 ± 0.833	(39) 11.39 ± 1.319	0.0022
3	(40) 107.4 ± 4.737	(40) 104.5 ± 4.784	0.6759	(41) 17.93 ± 1.287	(39) 21.62 ± 1.536	0.0684
4	(40) 16.55 ± 2.777	(40) 17.50 ± 2.728	0.8078	(41) 18.12 ± 1.446	(39) 17.26 ± 1.523	0.6811
5	(40) 1.700 ± 0.778	(40) 2.125 ± 0.875	0.7176	(40) 10.48 ± 1.332	(39) 10.67 ± 1.163	0.9141
6	(40) 0.700 ± 0.266	(39) 1.282 ± 0.593	0.3696	(39) 5.769 ± 0.944	(37) 5.919 ± 0.925	0.9102
7	(40) 0.300 ± 0.144	(37) 0.324 ± 0.169	0.9127	(35) 1.714 ± 0.421	(36) 3.250 ± 0.661	0.0556
8	(40) 0.050 ± 0.035	(34) 0.176 ± 0.130	0.317	(32) 1.500 ± 0.433	(32) 1.406 ± 0.411	0.8758
9	(39) 0.000 ± 0.000	(33) 0.030 ± 0.030	0.2801	(25) 0.880 ± 0.426	(30) 1.333 ± 0.478	0.49
10	(38) 0.237 ± 0.133	(32) 0.000 ± 0.000	0.1072	(20) 0.100 ± 0.100	(27) 0.259 ± 0.182	0.4893
11	(34) 0.000 ± 0.000	(32) 0.000 ± 0.000	.	(18) 0.000 ± 0.000	(27) 0.185 ± 0.131	0.2571
12	(34) 0.000 ± 0.000	(30) 0.000 ± 0.000	.	(14) 0.000 ± 0.000	(24) 0.000 ± 0.000	.
13	(33) 0.000 ± 0.000	(29) 0.000 ± 0.000	.	(8) 0.000 ± 0.000	(22) 0.000 ± 0.000	.
14	(29) 0.000 ± 0.000	(22) 0.000 ± 0.000	.	(5) 0.000 ± 0.000	(15) 0.133 ± 0.133	0.5778
Total	(40) 310.1 ± 5.577	(40) 307.5 ± 9.386	0.8142	(39) 61.46 ± 4.661	(41) 72.79 ± 5.440	0.1165

All values in parentheses are sample sizes; others are means ± 1 S.E.M. The last column shows p values associated with two-sided t-tests comparing N2 and N2^{MA431} production and gas-1 G0 production with gas-1^{MA431} production.

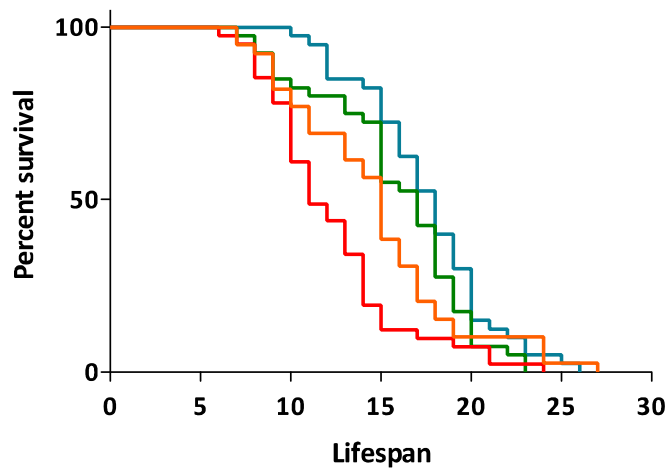


Figure 2.6. Age-specific survival for MA431 SNP-containing lines and ancestral controls. Percent daily survival for N2 (blue), N2^{MA431} (green), *gas-1* G0 (red), *gas-1*^{MA431} (orange).

MA431 SNP characterization: ROS and ATP. ROS levels differed among the four experimental strains (Figure 2.7; $F_{(3, 166)} = 20.95$, $p < 0.0001$), such that *gas-1*^{MA431} ROS was reduced compared to *gas-1* G0, and both *gas-1* (*fc21*)-containing strains exhibited higher ROS levels than N2 and N2^{MA431} at $p < 0.01$ (Tukey's HSD, $\alpha = 0.05$). Strains also varied in their ATP content (Figure 2.8; $F_{(3, 16)} = 9.781$, $p = 0.0007$), such that N2 had significantly higher ATP levels than either of the *gas-1* (*fc21*)-containing strains at $p < 0.01$, and marginally higher levels than N2^{MA431} at $p = 0.0695$ (Tukey's HSD, $\alpha = 0.05$). N2^{MA431}, *gas-1* G0 and *gas-1*^{MA431} levels were statistically indistinguishable.

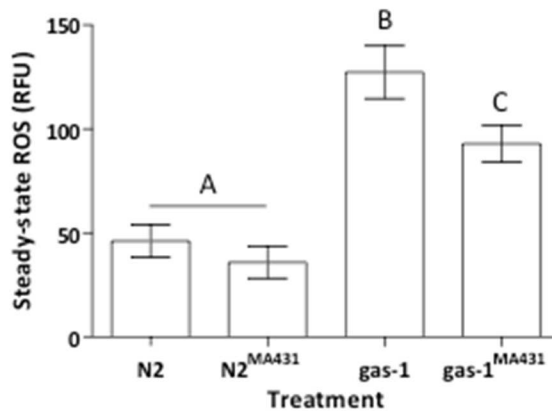


Figure 2.7. Steady-state ROS levels for MA431 SNP-containing lines and ancestral controls. ROS levels (mean \pm 1 S.E.M.) for N2 (46.17 ± 7.775), N2^{MA431} (35.95 ± 7.788), *gas-1* G0 (127.3 ± 12.79), and *gas-1*^{MA431} (92.94 ± 8.805). Error bars = 1 S.E.M. Letters group statistically indistinguishable groups.

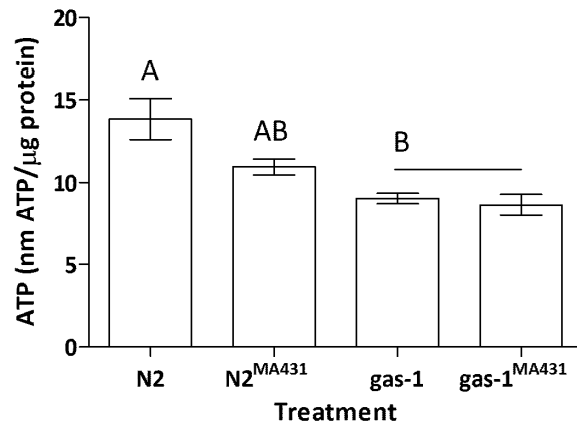


Figure 2.8. Steady-state ATP levels for MA431 SNP-containing lines and ancestral controls. ATP levels (mean ± 1 S.E.M.) for N2 (13.85 ± 1.242), N2^{MA431} (10.96 ± 0.480), *gas-1* G0 (9.042 ± 0.323), and *gas-1*^{MA431} (8.637 ± 0.661). Letters group statistically indistinguishable groups.

Discussion

In this chapter, I built upon previous work that combined the MA approach and high-throughput sequencing applied to two genetic backgrounds—one wildtype and one *gas-1* ETC mutant. The goals of the previous work were to assess the effect of increased endogenous ROS produced by the *gas-1* ETC mutant on MA line mutation rates and molecular spectra and on MA line fitness and phenotype. The MA process was expected to result in further fitness decline due to accumulated deleterious mutations (Kondrashov and Kondrashov 2010); however, the *gas-1* MA lines experienced partial fitness and phenotypic recovery towards wildtype levels (Lue 2015). We previously predicted an increased rate of new mutations alongside increasing endogenous steady-state ROS levels compared to wildtype. However, possibly owing to a recovery of steady-state ROS levels towards wildtype during MA, mutation rates in the *gas-1* MA lines were not significantly different than wildtype (Wernick 2016). The fitness and phenotypic recovery of the *gas-1* MA lines suggested that beneficial and/or compensatory mutations had accumulated in

the *gas-1* MA lineages; however, the fitness and phenotypic effects of individual mutations had not yet been characterized.

I added to this body of work by assessing the effect of the *gas-1* ETC mutation on the evolution of phenotypic means and variances during MA. I also characterized candidate mutations and pathways for mitigating ETC complex I impairment and identified patterns of selection in *gas-1* MA recovery. I determined that nuclear mutations were beneficial for fitness, compensatory for ROS, and context dependent for ATP. I also concluded that selection—both purifying and positive—can be a powerful force leading to maintenance or partial recovery of fitness in *gas-1* ETC-mutant MA lines, even under unfavorable population-genetic conditions. This is the first *C. elegans* MA study that functionally characterized putatively beneficial alleles that arose in an MA experiment that purportedly mitigate mitochondrial ETC dysfunction.

Evolution of gas-1 MA line phenotypes. In contrast to expectations and results of standard MA studies initiated from near optimally fit wild-type genotypes, we found that after fewer than 50 generations of single-individual bottlenecks, mean absolute and relative fitness of *gas-1* MA lines had improved slightly over that of their ancestral control with no marked increases in among-line variance of fitness-related phenotypes (Table 2-1, Table 2-2, Table 2-3). The reason for these departures from expectations are almost certainly related to the extremely low fitness of the starting *gas-1* (*fc21*) genotype. First, the observed increase in the number of nonviable eggs in MA lines compared to the *gas-1* G0 ancestor is indicative of purifying selection, and suggests that a significant number of new mutations in the *gas-1* MA lines were lethal. In addition, the large number of MA

line extinctions compared to the Baer N2 MA experiment provides further evidence for purifying selection at zygotic or postzygotic levels. Similarly, both the per generation change in among-line genetic variance, V_b , and mutational variance, V_M , for *gas-1* MA absolute fitness (W) were smaller than those estimated for the Baer N2 MA lines (Baer et al. 2005): $V_b = 5.191 \pm 1.824$ vs. 11.54 ± 2.64 and $V_M = -1.547$ vs. 5.77 , respectively. Second, there may be many evolutionary paths to (minor) fitness improvement available to the low-fitness *gas-1* genotype (c.f., (Estes and Lynch 2003; Burch and Chao 2004; McBride et al. 2008; Barrick et al. 2010. These results are in agreement with Fisher's adaptive landscape theory (Fisher 1930; Orr et al. 2003), and suggest that the fraction of mutations that are beneficial or compensatory increases with a population's distance from a fitness optimum. More specifically, our results are consistent with a scenario in which the action of purifying selection against sublethal and/or lethal mutations and positive selection favoring new (adequately supplied) beneficial or compensatory mutations in *gas-1* MA lines was sufficient to offset the negative impact of accumulated deleterious mutations, thereby maintaining or slightly improving fitness of individual *gas-1* lineages. Note that a previous MA experiment initiated from another ETC-deficient genotype, *mev-1*, with only slightly reduced fitness compared to N2, revealed the expected decline in mean fitness with successive MA generations (Joyner-Matos et al. 2011)—indicating that our results are not specific to ETC-deficient genotypes. Additionally, our findings suggest that, at low-fitness extremes, the DFE may shift to simultaneously offer an increased number of mutational routes to either fitness maintenance/recovery or extinction. That we observed only minor fitness improvements—nothing approaching complete recovery of ancestral

fitness levels—in *gas-1* MA lines could be due to a lack of medium- or large-effect beneficial mutations available to compensate for the effects of *gas-1* (*fc21*), or to their individual effects being muted by deleterious mutations also present in MA line genomes. The observation of MA line fitness improvement is in some ways reminiscent of results from MA experiments conducted using *Arabidopsis thaliana* by Shaw and colleagues (Shaw et al. 2002; Rutter et al. 2010) who found that, while among-line variance increased, mean fitness remained the same. One potential explanation for this result was cell lineage selection in the meristem; where somatic cells continuously undergo efficient selection before transitioning to germ cells.

In contrast to positive ΔM for *gas-1* MA absolute fitness and fitness relative to *gas-1* GO, MA had little impact on lifespan. This is consistent with previous studies in wildtype (Vassilieva et al. 2000) and ETC mutant *mev-1* (Joyner-Matos et al. 2011) populations and supports previous findings showing that V_m for lifespan needs many hundreds of generations of mutation accumulation to be affected.

Interactome and gene ontology analyses point to adaptation via altered membrane components and transport processes. In agreement with the phenotypic results, our bioinformatics analyses revealed less evidence for drift in *gas-1* MA line molecular evolution compared to N2 MA lines, and were consistent with the idea that some *gas-1* MA line mutations conferred beneficial or compensatory effects. First, fewer GO slim categories were significantly enriched in the *gas-1* MA line mutated gene set (two) compared to that of the N2 MA lines (twelve) (Figure 2.3). The GoMiner enrichment algorithm takes the total number of altered genes into account when calculating

enrichment scores, so this pattern is not a simple consequence of N2 MA lines having more mutations; it is instead consistent with genetic drift playing a weaker role in *gas-1* MA versus N2 MA line molecular evolution. Second, four genic mutations in the *gas-1* MA lines—*daf-2*, *rheb-1*, *twk-31*, and *smf-3*—had membership in the same GO slim categories—plasma membrane, transport, and cell (not significantly enriched)—suggesting some overlap in their function. It may be that altering the transport of nutrients, signaling molecules, or other proteins and ions across membranes might counterbalance *gas-1*-associated mitochondrial dysfunction. Given these genes' membership in important metabolic pathways (IGF, UPR^{mt}, K⁺ transport), it is possible that any ameliorating effects act by altering flux in these pathways at plasma membrane junctions. Third, that the process of embryonic development was significantly enriched only in the N2 MA lines suggests that mutations affecting this process were only mildly deleterious for wildtype N2, but potentially lethal and selectively eliminated within unfit *gas-1* lines. More specifically, a smaller fraction of *gas-1* MA line genic mutations (4/23, 17%) annotated to embryonic development compared to N2 MA mutations (16/55, 29%), with the latter exactly matching the fraction of all *C. elegans* genes annotating to the embryonic development ontology (3264/11293, 29%) (Appendices A and B). Thus, mutations in embryonic development genes may have a range of deleterious fitness effects depending on their genetic context. Fourth, the number of *gas-1* MA line mutations occurring within *gas-1* interactome genes was higher than expected by chance. This was not the case for the N2 MA line mutations and suggested that some *gas-1* MA line mutations were fixed in response to the *gas-1* mutation rather than resulting from

drift. And finally, four of the mutated genes that were members of the two significantly enriched *gas-1* MA GO slim categories (plasma membrane and transport) were also members of the *gas-1* interactome—*daf-2*, *rheb-1*, *smf-3*, and *sel-2*. Interestingly, despite three of these mutations having membership in the plasma membrane GO slim category, this category was not significantly enriched when considering all *gas-1* interactome genes. Although plasma membrane components do not comprise a significant portion of the *gas-1* interactome, they may have been an important target for compensatory or beneficial mutation in the *gas-1* MA lines.

Characterization of beneficial haplotype. We were unable to isolate the four candidate beneficial/compensatory *gas-1* MA431 SNPs for individual characterization. For the three chromosome III SNPs, this failure was quite likely due to their strong linkage; failure to isolate the chromosome V SNP from this group, however, may have been due to a beneficial epistatic interaction between this SNP and one or more of the others on chromosome III. We found no evidence that this pattern resulted from a translocation event involving chromosomes V and III; a caveat is that if a segment of highly repetitive DNA joined the translocated chromosome V to chromosome III, then the de novo assembly would be unable to construct a contig or scaffold across that region. In any case, we could not ascertain which SNP may have been the primary contributor to the slight *gas-1* MA431 fitness and phenotypic recovery.

Interestingly, the chromosome III/chromosome V SNP combination proved to be even more advantageous on the N2 background (Figure 2.4) rather than appearing to specifically compensate for the adverse effects of *gas-1* (*fc21*). This outcome was

somewhat surprising given the reasonable expectation that the N2 strain has achieved a local fitness optima under a constant laboratory environment (Gillespie 1984), and suggests a rugged genetic landscape in which a newly accessible evolutionary pathway was opened to N2 (Poelwijk et al. 2007), allowing the strain to achieve a new optimum. Two starkly different genetic explanations for this outcome include one of strong positive epistasis operating between the chromosome V SNP and some combination of those on chromosome III and the N2 background; in this scenario, the SNPs would be individually neutral or beneficial, but even more advantageous in combination (Phillips 2008). A second possibility is one of reciprocal sign epistasis (Kvitek and Sherlock 2011; Poelwijk et al. 2011) in which individually deleterious mutations confer a beneficial effect when in combination. As previously noted (Methods), when any one of four MA431 SNPs was lost during the backcrossing phase, all others would be lost in the following generation—an outcome consistent with reciprocal sign epistasis operating between some combination of the chromosome III and V SNPs. Sign epistasis has long been implicated in rugged fitness landscape (Wright 1932); however, only recently have laboratory studies observed the effect of sign epistasis on fitness optima in experimental evolution studies (Poelwijk et al. 2011) including in bacteria (*Comamonas sp.*) evolving in semi-novel environments (Korona et al. 1994), in the antibiotic enzyme TEM-1 (Salverda 2008), (Kvitek and Sherlock 2011).

Despite similarities in the effect of the MA431 SNPs on fitness in *gas-1* (*fc21*) versus N2, their effect at the physiological level were strongly strain-dependent; i.e., reduced steady-state ROS levels with no effect on steady-state ATP levels in *gas-1*, slightly

reduced steady-state ATP levels with no effect on steady-state ROS levels in N2. Thus, we can conclude the effect of the MA431 SNPs to be compensating for ROS on the *gas-1* background—an outcome consistent with population genetic predictions for evolutionary paths to mitigating the effect of mutations in an unfit population (Poon and Otto 2000). The observation of no effect of the MA431 SNPs on steady-state ATP (no change from *gas-1* G0) further suggests that continued decrease in ATP levels is beneficial on a *gas-1* background and/or that the deleterious effect of new mutation decreases the further a population exists from its fitness optimum (Eyre-Walker and Keightley 2007).

Conclusion. The evolutionary fitness trajectory of populations clearly depends on their location on the fitness landscape (Wright 1932; Orr 2000; Poon and Otto 2000; Orr et al. 2003). Unlike traditional MA experiments with wildtype populations, fitness in the ETC-mutant *gas-1* MA lines improved with little increase in among-line variance combined with evidence for zygotic or post zygotic selection suggesting that selection—both purifying and positive—can be an extremely powerful force even in conditions of extreme genetic drift. Functional characterization of a putatively beneficial haplotype that arose in an MA experiment revealed successful identification of specific genetic changes can lead to a partial recovery of fitness and phenotype in an ETC-deficient mutant strain of *C. elegans*. My finding that the *gas-1*^{MA431} beneficial haplotype was associated with a decrease in steady-state endogenous ROS is consistent with decreased ROS levels in *gas-1* MA in general (Lue 2015). Furthermore, my finding that steady-state ATP levels in the *gas-1*^{MA431} decreased compared to wildtype suggests that fermentation may be metabolic strategy to cope with increase oxidative stress (Lue 2015). However, it is also possible that

a decline in ROS levels is due to an increase stress response or altered transcriptional activity brought on by mutation accumulation as our characterization of MA431 SNPs might suggest.

CHAPTER 3: FITNESS RECOVERY IN LARGE POPULATIONS

Background

Mutation is the ultimate source of genetic variation in populations. However, most new mutations are deleterious for organismal fitness and can propagate within populations under a broad range of conditions (Crow and Kimura 1970; Eyre-Walker and Keightley 2007; Kondrashov and Kondrashov 2010). When multiplied across the total number of sites within a genome, the cumulative effects of mildly deleterious mutation accumulation can add up to a large impact. If such mutations are not purged by selection, they can lead to a steady decline in fitness over evolutionary time—perhaps posing a special threat to the survival and integrity of small populations (Lande 1994; Lynch, J. Conery, and Burger 1995).

The fact that populations persist at high levels of fitness under this constant pressure of deleterious mutations suggests that other evolutionary processes can ameliorate the phenotypic decline caused by deleterious mutations (e.g., Estes et al. 2011). Three mutational processes able to counteract deleterious mutation accumulation include: 1) reversion mutations that restore wildtype sequences, 2) generally beneficial mutations that improve organismal fitness irrespective of genetic background, and 3) compensatory mutations that specifically mitigate the effects of previously-acquired deleterious mutations through epistasis (Kimura 1985). Unlike generally beneficial mutations, compensatory mutations are beneficial *only* within the context of a particular deleterious mutation and are, by themselves, either neutral or mildly deleterious. The potential for any of these mutation types alters our expectations for the impact of

deleterious mutation in populations or (Orr 2000; Poon and Otto 2000a; Lanfear et al. 2014), but since the fitness effects of individual mutations are rarely characterized, the relative importance of beneficial and compensatory epistatic mutations is unknown.

The rate and repeatability of beneficial or compensatory mutations fixing in a population will depend on the rate of new mutation, mutational selective effects (s), and by the population's effective size (N_e). As previously described in the background of Chapter 2, new mutations' rate and distribution of effects of fitness plays an important role in small population survival. However, as N_e increases, selection becomes more efficient at preventing deleterious mutation accumulation and promoting fixation of adaptive mutations; i.e., those with $s > 1/2N_e$ (Crow and Kimura 1970; Ohta 1992). Large sexual populations also benefit from genetic recombination that can facilitate the spread of adaptive alleles (Smith 1978) or combine alleles that together confer an adaptation (Birdsell and Willis 2003).

The population level consequences of adaptation through compensatory mutation have been well studied in bacteriophage (Rokyta et al 2002; Poon and Chao 2005; Rokyta et al. 2009), yeast (Kenigsberg et al. 2010; Szamecz et al. 2014), viruses and microbes (Maisnier and Anderson 2004; Gonzalez-Ortega et al. 2011; Springman et al. 2012), and *C. elegans* (Estes and Lynch 2003; Chandler et al 2011; Estes et al. 2011). These studies have found that although deleterious mutations are common in populations, they have the potential to partially mitigate their effects through fixation of compensatory mutations. Therefore, the potential for compensatory epistatic or generally beneficial mutations further alters our expectations for the impact of deleterious mutation in a

population (Orr 2000; Poon and Otto 2000). However, the capacity of adaptive evolution to restore ancestral phenotypes and the molecular genetic patterns of recovery are still poorly understood in complex eukaryotic systems.

The mitochondrial electron transport chain (ETC) offers a compelling framework for studying the accumulation and effects of mutations. Mitochondria are essential eukaryotic organelles that generate the majority of cellular bioenergy (ATP), but also produce ROS (Murphy 2009). ETC components pump protons across the mitochondrial intermembrane creating a proton motive force, which drives the production of ATP. Because the mitochondrial ETC is highly conserved, vital for energy metabolism in all eukaryotic life, it is an ideal system in which to study the impact of deleterious mutation. Many features of mitochondrial biology—high mutation rate, apparent lack of recombination, uniparental inheritance, and proximity to potentially damaging ROS—could potentially promote faster rates of evolution in mtDNA (Ballard and Whitlock 2004). In addition, since proper ETC functioning relies on the maintenance of favorable mitonuclear epistatic interactions (Blier 2001; Dowling et al. 2008), it is an excellent system for addressing the role of epistasis in maintaining population fitness. The capacity of adaptive mutation to recover mitochondrial dysfunction remains largely unstudied.

In this chapter, I describe adaptive evolution of the mitochondrial ETC in real-time using the ETC-deficient *gas-1* *Caenorhabditis elegans* nematode strain compared to the wildtype Bristol N2 strain. The *gas-1* strain contains a single missense mutation affecting ETC Complex I (Kayser et al. 1999) and exhibits deleterious phenotypes including: low reproductive fitness, high levels of potentially damaging reactive oxygen species (ROS),

and low ATP levels, but slow aging, relative to wildtype (Hartman et al. 2001; Kayser et al. 2004; Lenaz et al. 2006; Lue 2015). The goal of this research was to understand the capacity of new mutations to recover ancestral (wildtype) levels of fitness and phenotype of the *gas-1* strain experimentally evolved under efficient selection conditions (ie. large populations).

I examine genetic patterns and phenotypic mechanisms of fitness recovery in 24 replicate lineages of the *gas-1* strain (G0) experimentally evolved by Jennifer Sullins (S. Estes Lab, Portland State University) for 60 generations in non-overlapping, large populations ($n = 1000$). I assessed fitness with a conventional fitness assay and compared it to a competitive fitness assay performed by the L. T. Morran Lab (Emory University, Atlanta, GA). I found that average fitness of evolved lines increased as expected, but that two distinct groups of lines emerged—one that evolved significantly higher average fitness, and containing some lines that either recovered or exceeded wildtype fitness levels, and another group that showed more modest fitness gains. Interestingly, many lines in the first group were observed to generate significant numbers of males during experimental evolution, consistent with evolution of outcrossing either accompanying or driving rapid fitness recovery (Morran et al. 2009). To statistically assess the presence of males in the *gas-1* RC lines and explain their mode of generation, I performed male frequency and non-disjunction assays on a subset of the *gas-1* RC lines observed as generating males, a random sampling of the other *gas-1* RC lines, *gas-1* G0, and wildtype N2. I then bioinformatically characterize the nuclear mutations that arose in the *gas-1* RC lines during experimental evolution and describe patterns in their functional

characteristics. My research advances knowledge about the capacity of new, potentially beneficial or compensatory mutations to mitigate ETC dysfunction and restore ancestral phenotypes in large populations. This research also broadens our current knowledge base on the genetic patterns and mechanisms of putatively compensatory or beneficial mutations in mitochondrial dysfunctional populations.

Methods

Strains and culture conditions. This thesis utilized recovery (RC) lines generated from a *gas-1* mutant. The *gas-1* mutant, derived from ethyl methanesulfonate (EMS) mutagenesis (Kayser et al. 1999), was obtained from the Caenorhabditis Genetics Center (University of Minnesota) and repeatedly backcrossed to the standard laboratory wildtype strain, Bristol N2, in an attempt to create an isogenic mutant strain. As detailed in Wernick et al. 2016, a *gas-1* mutant strain, CW152, was backcrossed to Bristol N2 for 10 generations during which time PCR and Sanger DNA sequencing were used to track the *gas-1* (*fc21*) allele, a C → T point mutation that replaces a conserved arginine with lysine in the GAS-1 protein. Jennifer Sullins (S. Estes Lab at Portland State University) generated the 24 *gas-1* RC lines using improved upon methods of (Estes and Lynch 2003) and (Denver et al. 2010). The *gas-1* strain was allowed to recover from freezing for two generations at 20 °C on Nematode Growth Medium-Lite plates seeded with *Escherichia coli* OP50-1. After recovery, 24 randomly chosen L4-stage offspring from a single *gas-1* G0 parent were used to initiate 24 RC lines. The 24 replicate lineages of *gas-1* were contemporaneously passaged for 60 non-overlapping generations. Approximately 1000 individuals were transferred each generation. Population sizes were estimated via sample

counts and non-overlapping generations were maintained by standard age synchronization (Phillips et al. 2014) of cultures prior to each transfer. This large-population size treatment minimizes the effect of genetic drift and promotes efficient selection against deleterious mutations and in favor of beneficial mutations, and the 60-generation timescale was sufficient for fitness recovery in other mutant strains (Estes et al. 2011a). Allowing the replicates to evolve under non-overlapping conditions simplifies our statistical analysis and sequencing interpretation, while acting as a good approximation for populations with more complex life histories (Hartl and Clark 2007). We froze samples of the RC lines on the fourth and fifth generations and every five generations thereafter until generation 60 as back-up.

Prior to all phenotypic experiments, strains were allowed to recover from freezing for two to three generations at 20 °C on Nematode Growth Medium-Lite (NGML) plates containing 20 µg/mL streptomycin seeded with *Escherichia coli* OP50-1. For the ROS and non-disjunction assays, lines were age synchronized by standard bleach treatment prior to each assay.

Conventional fitness assay. I assayed daily offspring production and lifespan on N2, *gas-1* G0 and the *gas-1* RC lines (Gmax 60) following established methods (Estes et al. 2003; Estes; Phillips, et al. 2011; and see Chapter Two). Prior to the assay, lines were segregated into ten replicates for N2 and *gas-1* G0 and five replicates for each *gas-1* RC line during recovery from freezing and one adult hermaphrodite (“parent” below) from each replicate was transferred to an individual plate to lay eggs for five hours, creating an age-synchronous cohort. One offspring from each parent was then transferred to individual

seeded plates, resulting in a total of 3-5 individuals from each *gas-1* RC line and 9-10 control individuals assayed.

Individuals were transferred daily until cessation of egg laying to separate parents from offspring. Plates were incubated at 20 °C for 36 hours to allow eggs to hatch; resulting offspring were killed with one drop of 0.5 M sodium azide and counted directly by contrast staining plates with toluidine blue. I also recorded the number of unhatched and presumably unviable eggs, which are easily distinguished from unfertilized eggs under a light microscope. Lifespan was calculated as days lived from the egg stage; survival rates were calculated from the same dataset. Eight worms died due to desiccation after crawling up the side of Petri plates and were omitted from analysis.

The above data were used to calculate mean relative fitness (ω) of the *gas-1* G0 mutant compared to N2 and of each *gas-1* RC line compared to its *gas-1* G0 ancestor. Relative fitness of each individual was computed following (Peters et al. 2003) as: $\omega = \Sigma e^{-rx} l(x) m(x)$, where $l(x)$ is the number of worms surviving to day x , $m(x)$ is the fecundity at day x , and r is the mean intrinsic population growth rate of the assay-specific N2 or *gas-1* control as appropriate. The latter was calculated by solving Euler's equation for r from the equation $\omega = \Sigma e^{-rx} l(x) m(x) = 1$ using an average value of $l(x) m(x)$ for the controls. I set $x = 4.75$ on the first reproductive day (cf. Vassilieva et al. 2000).

Competitive fitness assay. In conjunction with the conventional fitness assay, which revealed two classes of fitness recovery—lower and higher—among the *gas-1* RC lines compared to *gas-1* G0, the L. T. Morran Lab at Emory University (Atlanta, GA) performed a competitive fitness assay on a subset of the RC lines that represented a mix of lines from

both lower and higher fitness classes—RC2, RC3, RC6, RC8, RC9, RC11, RC12, RC16, RC17, RC18, RC22, RC23, RC24—and the *gas-1* G0 progenitor. The competitive fitness assay measures fitness by competing lines against a GFP-marked (green fluorescence protein) tester strain (*JK2735*) following the methods of Morran & Phillips, 2009. These assays more closely simulate the environment in which the worms were experimentally evolved—where food eventually becomes limiting. We can thus evaluate fitness in the same environment in which mutations have accumulated (i.e. one with strong competition for food resources) and thereby avoid potential genotype-by-environment interactions. Competitive fitness assays were conducted in two blocks to obtain a total of 2-8 replicates of recovery lines (with a goal of five) and on five replicates of *gas-1* G0.

To begin each competition, roughly equal numbers of age-synchronized L4 RC individuals were mixed with age-synchronized L4 GFP-marked individuals on seeded plates. The goal was to plate 500 of each, though this number ranged from 402 – 666 individuals. The RC lines or *gas-1* G0 strain, and GFP individuals were allowed to mingle and mate. After approximately three days (i.e., one generation) L4 offspring from the matings were counted. The GFP-expression in the GFP-marked strain is dominant; therefore, if any outcrossing occurred, then fitness would be underestimated for the experimental lines or *gas-1* G0 strain. (Care was therefore taken to ensure no males were added to the competition assays.) Prior to counting, worms were washed in M9 containing 10-20% EtOH, killing all worms, onto unseeded plates. Only L4 offspring were counted. The ratio of non-GFP-marked RC or *gas-1* G0 offspring to GFP-marked offspring

was calculated for each replicate. These ratios were used to gauge competitive fitness of the RC lines and the *gas-1* G0 strain relative to each other.

Life-history data analysis. I first analyzed differences in life history trait means between *gas-1* G0 and N2 using *t* tests. We also assessed initial life history trait within-line variances of *gas-1* G0 and N2 using one-way ANOVAs. Next, I tested for life history mean differences between *gas-1* RC and *gas-1* G0 using two-way mixed-model ANOVAs with the formula $y = \text{strain} + \text{line}(\text{strain}) + \varepsilon$ with *strain*—*gas-1* G0 or *gas-1* RC—as a fixed effect and *line*(*strain*)—each RC line nested within *gas-1* RC—as a random effect using the expected mean squares (EMS) method as applied in JMP12 (SAS); *line*(*strain*) for *gas-1* G0 has no effect on the model. I then tested pairwise differences in means between *gas-1* G0 and all *gas-1* RC lines individually using post hoc Dunnett’s tests based off of within the two-way mixed model ANOVAs above. I used the Cox proportional hazards method to compare age-specific mortality of N2 and *gas-1* G0, along with the *gas-1* RC lines as a group, individually, and broken down by male-containing lines, with *strain* and *line*(*strain*) as main effects when assessing the RC lines individually. For the RC lines, within- and among-line components of variance were calculated using restricted maximum likelihood (REML) as applied in JMP12 (SAS) while restricting variance estimates to be positive. The one-way ANOVA model for analyzing the *gas-1* RC lines was $y = \mu + \text{RC line} + \varepsilon$, wherein *RC line* is a random effect and represents among-line variance and ε represents within-line component of variance. Finally, I performed linear regression analyses on the number of SNPs and SNP types (i.e., exonic, intronic, etc.) and life history traits to determine relationships between mutation and life-histories; when assessing SNP relationships to

the presence of males and bagging, I used contingency tables with continuity correction. Correlation between competitive fitness and relative fitness—both to N2 and *gas-1* G0—was assessed using a Spearman's rank coefficient and least squares regression. All statistical tests were performed in JMP12 (SAS).

Male frequency and non-disjunction assays. Along with undergraduate student Joseph Ramirez, I performed male frequency assays on all *gas-1* RC lines that began generating observable numbers of males in the 60-generation evolution experiment and on *gas-1* RC lines representing a mix of lower and higher fitness classes—RC2, RC3, RC6, RC7, RC8, RC9, RC10, RC11, RC12, RC15, RC16, RC17, RC18, RC22, RC23, and RC24—along with *gas-1* G0 and N2 wildtype. Males and hermaphrodites were counted on four replicate plates containing roughly 200 age-synchronized young adult worms; the number of worms on each plate ranged from 53-429, however. Average male to hermaphrodite proportions were then calculated for each *gas-1* RC line assayed along with *gas-1* G0 and wildtype N2. Non-disjunction assays were performed on the same *gas-1* RC lines as the male frequency assays along with *gas-1* G0 and N2 wildtype. For the non-disjunction assays, we counted males and hermaphrodites on six replicate plates containing roughly 200 (range: 107-612) age-synchronized young adult offspring of 10-20 unmated (virgin) hermaphrodites. Average male to hermaphrodite proportions were then calculated for each *gas-1* RC line assayed along with *gas-1* G0 and wildtype N2.

Identification and characterization of *gas-1* RC line mutations; evolutionary rate analysis. Riana Wernick in the (D. R. Denver Lab at Oregon State University (Corvallis, OR)) whole-genome sequenced (Illumina DNaseq) the 24 *gas-1* RC lines. Samples were

prepared from first larval (L1) stage nematodes as described in (Wernick et al. 2016). This procedure was previously performed for the *gas-1* G0 progenitor and N2 wildtype control in Wernick et al. 2016. Evolutionary rate was calculated from pooled *gas-1* RC lines with the equation $K = m / (LnT)$ where K is the base substitution rate (per nucleotide site per generation), L is the number of RC lines, m is the number of observed mutations, n is the number of nucleotide sites, and T is the time in generations as previously described (Denver et al. 2009). Candidate SNPs in *gas-1* RC lines were identified as variants from the *C. elegans* reference genome (WS242) and our wildtype N2 lab strain as previously described in (Wernick 2016; and see **Error! Reference source not found.**).

Gene ontology enrichments. I examined gene ontology (GO) enrichments for all nuclear DNA SNPs that arose within protein coding regions of *gas-1* RC lines within three functional domains—biological process, cellular component, and molecular function. I used GoMiner (Zeeberg et al. 2003, application build 457) and the GO MySQL Database (MySQL 3.7.13, Oracle Corporation, Cupertino, CA; GO database build 2016-06-07, geneontology.org) to calculate GO enrichments. GoMiner reports a two-sided Fisher’s exact p value based upon the number of genes in a category and a false discovery rate (FDR)-corrected p value for multiple testing (Zeeberg et al. 2005); however, the FDR can be overly conservative when the sample size or number of discoveries is small; i.e., less than 100 (Tong and Zhao 2008; Zeeberg et al. 2005), so I based my interpretations on results of the Fisher’s exact tests with $\alpha = 0.05$. I then determined which broader-level functional “GO slim” categories, as defined by the GO Consortium (Version 1.2, 2012-09-21, geneontology.org), were enriched using CateGORizer (Hu et al. 2008). CateGORizer

classifies enriched GO terms into their respective GO slim categories, giving a coarser-scale overview of enrichment patterns. GO slim categories were then checked against the GoMiner output to assess statistical significance.

Results

Evolution of male frequency during recovery. J. Sullins noted approximately midway through the 60-generation evolution experiment, we noted that six of the 24 replicate recovery lines—RC3, RC8, RC9, RC12, RC18, and RC24—began generating appreciable numbers of males, atypical in this androdiecious species (Anderson et al. 2012; Teotonio et al. 2012). Furthermore, two of the 24 lines—RC20 and RC22—became resistant to bleaching during the age-synchronization step between generations, requiring longer bleach treatment than their peer replicates to harvest eggs for the next generation. We found no significant rates of non-disjunction in any of the *gas-1* RC lines or *gas-1* G0 compared to wildtype N2. Male frequency assays revealed that only *gas-1* RC9 and RC17 had significantly greater numbers of males than *gas-1* G0 (Dunnett's test, $\alpha = 0.05$); however, a major caveat to this finding is that the lines were frozen prior to assays and it is difficult to recover any observable number of males after recovery from freezing (L. T. Moran, pers. comm.). Thus, the true frequency of males present in the above *gas-1* RC lines during the experiment was likely higher than reflected by these assays; if true, we can be confident that increased male frequency was not due to increased rates of non-disjunction in recovery lines.

Life history mean comparisons of *gas-1* G0 to Bristol N2. As expected based on previous studies of the CGC *gas-1* (*fc21*) strain (Morgan and Sedensky 1994; Kayser et al. 1999; and see Chapter 2), our backcrossed *gas-1* G0 strain exhibited extremely low offspring production compared to wildtype N2 (**Figure 3.1**; Figure 3.2; Table 3.1), resulting in significantly lower relative fitness than what? ($t_{(17)} = 8.015$, $p < 0.0001$; **Figure 3.1**; Table 3.2). In further agreement with our previous findings (Chapter 2), *gas-1* G0 reproductive maturity was delayed by one day; however, here we find its peak reproductive output delayed by a two days (rather than one day) compared to N2 (**Figure 3.1**; Figure 3.2; Table 3.2). We found no significant differences in reproductive span, although *gas-1* G0 had an average maximum reproductive period lasting four days longer than N2 (**Table 3.3**). In contrast to previous reports of *gas-1* being short-lived (Kayser et al. 2004), our *gas-1* G0 lived slightly longer than N2 (nonsignificant) (Table 3.2); average lifespan for our *gas-1* G0 strain (15.20 days) was more than five days longer than that reported for the CGC *gas-1* (*fc21*) strain (9.8 days) (Kayser et al. 2004). There were no significant differences in survival between the two strains according to a Cox proportional hazards model (data not shown). Finally, there were no significant differences in bagging (where internal hatching of eggs kill the parent) or the number of unhatched but fertilized eggs between *gas-1* G0 and wildtype N2 (Table 3.2).

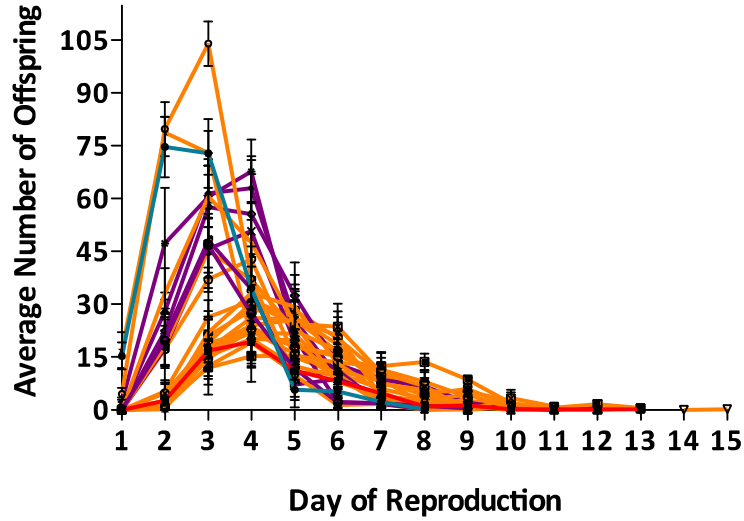


Figure 3.1. Reproductive schedules of individual *gas-1* RC lines, *gas-1* G0, and wildtype N2. Error bars are 1 S.E.M. Daily offspring production of N2 (blue, circles), *gas-1* G0 (red, squares), individual *gas-1* RC high-male lines (purple), and individual *gas-1* RC no male lines (orange). Summary statistics are reported in Table 3.2.

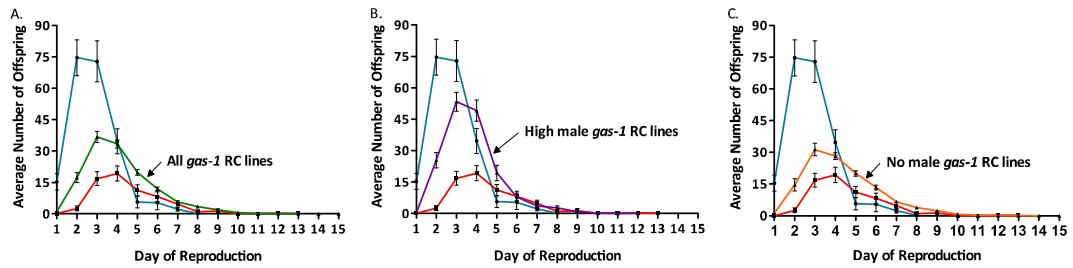


Figure 3.2. Reproductive schedules of *gas-1* RC, *gas-1* G0, and wildtype N2. Error bars are 1 S.E.M. A) Daily offspring production of N2 (blue, circles), *gas-1* G0 (red, squares), and the combined *gas-1* RC lines (green, triangles). Summary statistics are reported in Table 3.2. B) Daily offspring production of N2 (blue, circles), *gas-1* G0 (red, squares), and the *gas-1* RC lines containing males (purple, triangles pointing up). C) Daily offspring production of N2 (blue, circles), *gas-1* G0 (red, squares), and the *gas-1* RC lines without males (orange, triangles pointing down).

Table 3.1. One-way ANOVAs for life history traits for N2, *gas-1* G0, *gas-1* RC reported as means or variances \pm S.E.M.

Treatment	Mean	Among-line variance	Within-line variance
Absolute fitness (<i>W</i>)			
N2 Control	210.7 \pm 18.91		3217 \pm 1609
<i>gas-1</i> G0	64.20 \pm 11.97		1432 \pm 675.1
<i>gas-1</i> RC	131.0 \pm 6.278	2035 \pm 757.8	2468 \pm 369.6
Fitness relative to N2			
N2 Control	1.000 \pm 0.117		0.124 \pm 0.062
<i>gas-1</i> G0	0.095 \pm 0.021		0.004 \pm 0.002
<i>gas-1</i> RC	0.290 \pm 0.027	0.067 \pm 0.021	0.018 \pm 0.003
Fitness relative to <i>gas-1</i> G0			
N2 Control	7.026 \pm 0.736		4.871 \pm 2.436
<i>gas-1</i> G0	1.000 \pm 0.195		0.381 \pm 0.180
<i>gas-1</i> RC	2.584 \pm 0.196	3.316 \pm 1.045	1.079 \pm 0.162
Lifespan			
N2 Control	14.56 \pm 2.082		39.03 \pm 19.51
<i>gas-1</i> G0	15.20 \pm 1.306		17.07 \pm 8.045
<i>gas-1</i> RC	15.28 \pm 0.500	0	28.26 \pm 3.776

Table 3.2. Means for life history traits for N2, *gas-1* G0, and *gas-1* RC.

Character	N2	Controls			<i>gas-1</i> RC		<i>gas-1</i> RC by males	
		<i>gas-1</i> G0	t statistic	p value	As a group		Yes	No
Total Reproduction	(9) 210.7 ± 18.91	(10) 64.20 ± 11.97	6.688	< 0.0001	(113) 131.0 ± 6.278	(28) 162.9 ± 12.53	(85) 120.5 ± 6.921	
Lifespan	(9) 14.56 ± 2.082	(10) 15.20 ± 1.306	0.268	0.792	(113) 15.28 ± 0.500	(28) 17.11 ± 1.131	(85) 14.68 ± 0.539	
Fitness relative to N2	(9) 1.000 ± 0.117	(10) 0.095 ± 0.021	8.015	< 0.0001	(113) 0.290 ± 0.027	(28) 0.400 ± 0.039	(85) 0.254 ± 0.033	
Fitness relative to <i>gas-1</i> G0	(9) 7.026 ± 0.736	(10) 1.000 ± 0.195	8.305	< 0.0001	(113) 2.584 ± 0.196	(28) 3.541 ± 0.300	(85) 2.269 ± 0.232	
Unhatched eggs	(9) 1.222 ± 0.401	(10) 4.000 ± 1.333	1.903	0.0741	(113) 1.973 ± 0.193	(28) 1.571 ± 0.269	(85) 2.106 ± 0.240	
Bagging	3/9 (33.33%)	1 /10 (10.00%)	1.552	0.2129	44/113 (38.94%)	8/28 (28.57%)	36/85 (42.35%)	

All values in parentheses are sample sizes; others are means ± 1 S.E.M. Two-sided t-test statistics and associated p values are shown comparing *gas-1* G0 and N2; * denotes fraction of bagging individuals with percent in parentheses, X² test with X² statistic and continuity correction p value reported comparing *gas-1* G0 and N2.

Table 3.3. Summary statistics for daily reproduction for *gas-1* G0 and N2.

Reproductive Day	N2	<i>gas-1</i> G0	<i>t</i> statistic	<i>p</i> value
1	(9) 15.33 ± 3.830	(10) 0.000 ± 0.000	4.234	0.0006
2	(9) 74.67 ± 8.546	(10) 2.600 ± 1.157	8.818	< 0.0001
3	(9) 72.89 ± 9.770	(10) 16.80 ± 3.300	5.68	< 0.0001
4	(9) 34.67 ± 4.944	(10) 19.30 ± 4.691	2.255	0.0376
5	(9) 5.667 ± 2.851	(10) 11.20 ± 2.705	1.408	0.1771
6	(9) 5.333 ± 3.395	(9) 8.222 ± 2.368	0.698	0.4952
7	(8) 2.250 ± 0.921	(9) 4.778 ± 1.579	1.337	0.2011
8	(5) 0.200 ± 0.200	(9) 1.000 ± 0.882	0.659	0.522
9	(5) 0.000 ± 0.000	(9) 1.333 ± 1.106	0.883	0.3947
10	(5) 0.000 ± 0.000	(8) 0.125 ± 0.125	0.777	0.4533
11	(5) 0.000 ± 0.000	(7) 0.143 ± 0.143	0.833	0.4241
12	(5) 0.000 ± 0.000	(7) 0.286 ± 0.286	0.833	0.4241
13	(5) 0.000 ± 0.000	(6) 0.167 ± 0.167	0.905	0.3893

All values in parentheses are sample sizes; others are means ± 1 S.E.M. The last column shows *p* values associated with two-sided *t*-tests comparing *gas-1* G0 and N2.

Life-history evolution of *gas-1* RC lines. After 60 generations of evolution in large, non-overlapping populations, neither absolute fitness nor relative fitness of the *gas-1* RC lines improved significantly over *gas-1* G0 on average (Table 3.4). There was, however, significant variability in the responses of individual *gas-1* RC lines (Figure 3.1; Figure 3.3; Table 3.4), reflected in the increased phenotypic variance of the *gas-1* RC lines compared to *gas-1* G0 (Table 3.1). Lines RC3, RC8, RC9, RC10, RC13, RC20, and RC22 had significantly higher *W* than *gas-1* G0 (Dunnet's test, *p* < 0.05). The same lines, in addition to RC12 and RC18, had significantly higher fitness relative to *gas-1* G0; lines RC3, RC8, RC9, RC10, RC18, RC20, and RC22 had significantly higher fitness relative to wildtype N2 (Dunnet's test, *p* < 0.05). The *gas-1* RC lines showing significant improvements for absolute fitness and fitness relative to *gas-1* G0 and N2 largely overlap those we noted as generating appreciable numbers of males in the 60-generation evolution experiment (Figure 3.1; Figure 3.2B vs. 3.1C). Furthermore, *gas-1* RC20 and RC22, which became resistant to bleaching during

the age-synchronization step between generations are also among the lines showing significant fitness improvements. There were no significant differences, either between strains (*gas-1* RC vs. *gas-1* G0) or among *gas-1* RC lines, for lifespan (Table 3.1), survival (Cox proportional hazards, data not shown), the presence of unhatched eggs (Table 3.2), or the prevalence of bagging, even when the presence of males was taken into account (Table 3.2).

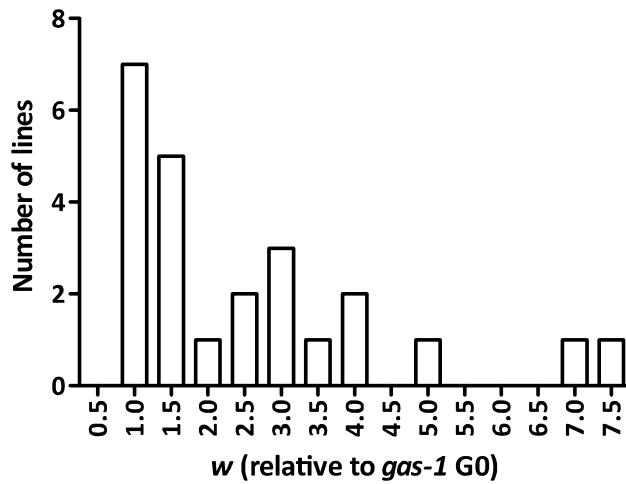


Figure 3.3. Frequency histogram of the 24 *gas-1* RC lines' fitness (w) relative to *gas-1* G0 (equal to 1.0).

Table 3.4. Two-way mixed-model ANOVA for life history traits in *gas-1* RC and *gas-1* G0. Statistical analyses are detailed in methods.

Fitness relative to N2					
Source	SS	d.f.	MS	F ratio	p value
Strain	0.332	1	0.332	0.489	0.4919
Line(Strain)	7.896	23	0.343	20.44	<0.0001
Model	8.245	24	0.344	20.45	<0.0001
Error	1.646	98	0.017		
Total	9.892	122			
Fitness relative to <i>gas-1</i> G0					
Source	SS	d.f.	MS	F ratio	p value
Strain	22.06	1	22.06	0.659	0.4257
Line(Strain)	390.8	23	16.99	16.73	<0.0001
Model	413.9	24	17.24	16.98	<0.0001
Error	99.51	98	1.015		
Total	513.4	122			
Absolute fitness					
Source	SS	d.f.	MS	F ratio	p value
Strain	39790	1	39790	1.793	0.1966
Line(Strain)	279000	23	12130	5.108	<0.0001
Model	320000	24	13330	5.615	<0.0001
Error	230000	98	2375		
Total	550000	122			
Lifespan (days)					
Source	SS	d.f.	MS	F ratio	p value
Strain	0.018	1	0.018	0.001	0.9804
Line(Strain)	630.5	23	27.41	0.999	0.4737
Model	630.5	24	26.27	0.958	0.5268
Error	2688	98	27.43		
Total	3319	122			

Comparison of relative fitness and competitive fitness. The *gas-1* G0 strain performed worse than most *gas-1* RC lines in competition against the GFP-marked strain, and none of the G60 *gas-1* RC lines outcompeted the GFP-marked strain, although eight of them competed at levels >70% of the GFP tester strain (Figure 3.4). Interestingly, RC22, which performed the best in terms of relative fitness (both to *gas-1* G0 and N2; data not shown)

ranked fifth in the competitive fitness assay; however, relative and competitive fitness of individual RC lines was positively correlated ($\rho = 0.5500$, $p = 0.0337$) (Figure 3.4). The *gas-1* RC lines noted as generating high frequencies of males tended to perform better in both types of fitness assays than those without males, but performed especially well in the competitive fitness assays (Figure 3.4, purple dots). With two noteworthy exceptions (RC7 and RC11), most lower-relative fitness lines cluster about the origin (Figure 3.4), suggesting that differences in performance in the two types of assays are only realized when the magnitude of fitness change is greatest.

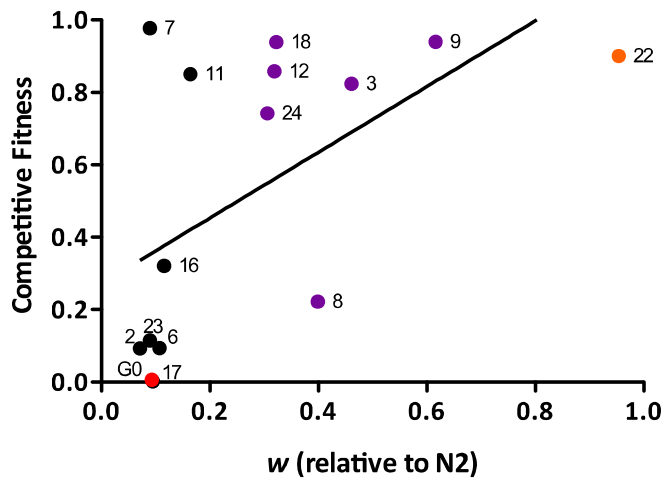


Figure 3.4. Correlation of relative fitness with competitive fitness. Data points represent *gas-1* RC lines along with *gas-1* G0 that were assayed for both fitness relative to N2 and competitive fitness to the GFP-marked strain. Black dots indicate *gas-1* RC lines without evidence for males, purple indicates *gas-1* RC lines in which the presence of males was noted, orange indicates the bleach resistant line that was included in the competitive fitness assay, and red indicates *gas-1* G0 which overlaps with *gas-1* RC line 17. Black line indicates best fit line of the data using standard least squares regression where y is competitive fitness and x is relative N2 fitness: $y = 0.908(x) + 0.271$. Best fit statistics ($F_{(1, 13)} = 6.083$, $p = 0.0283$), non-significant intercept.

Identification and characterization of *gas-1* RC line mutations. The Dee Denver Lab at Oregon State University (Corvallis, OR) whole-genome sequenced (Illumina DNaseq) the

gas-1 RC lines, along with *gas-1* G0 and wildtype N2, and identified and characterized the nuclear mutations. A total of 113 SNPs was detected; of these, 21 occurred in intron regions, 18 in exons, and the remaining 64 mutations were intergenic. Of the 18 exonic mutations, 7 were synonymous, and 11 were nonsynonymous. RC-line specific mutations numbers ranged from 1 (in RC14, RC15, and RC21) to 11 (in RC13). The evolutionary rate (substitution rate) in *gas-1* RC lines, $8.02 \times 10^{-10} \pm 7.54 \times 10^{-11}$, was 2.6 times slower than the evolutionary rate (mutation rate) in the *gas-1* MA lines of $2.12 \times 10^{-9} \pm 3.36 \times 10^{-10}$ per a site per a generation and 3.4 times slower than the evolutionary rate (mutation rate) in the Baer N2 MA lines of $2.7 \times 10^{-9} \pm 1.36 \times 10^{-10}$ (Denver et al. 2009). The evolutionary rate value of the *gas-1* RC lines and mutation rate of *gas-1* MA lines and N2 values did not overlap within three-standard errors of the mean, respectively.

GO term enrichment. Four GO slim categories were significantly enriched in the *gas-1* RC lines: the biological processes of 1) cell organization and biogenesis (GO:0016043), 2) cytoskeleton organization and biogenesis (GO:0007010), and 3) protein modification (GO:0006464), and the molecular function category of 4) phosphoprotein phosphatase activity (GO:0004721); no cellular component categories were significantly enriched (Figure 3.5). Ten *gas-1* RC-line mutated genes mapped to the cell organization and biogenesis term: *ttn-1*, *syd-2*, *slt-1*, *tir-1*, *fer-1*, *ppfr-1*, *math-33*, *ced-12*, *Y34B4A.2*, and *pph-4.1*; four SNPs mapped to cytoskeleton organization and biogenesis: *ttn-1*, *ppfr-1*, *ced-12*, and *pph-4.1*; nine SNPs mapped to protein modification: *ttn-1*, *tir-1*, *C17H12.3*, *Y54F10BM.3*, *ppfr-1*, *Y106G6D.4*, *math-33*, *pph-4.1*, and *set-29*; and three SNPs mapped to phosphoprotein phosphatase activity: *Y54F10BM.3*, *pph-4.1*, and *C17H12.3* (Appendix

D). The significantly enriched GO categories to which the largest numbers of SNPs mapped were the biological process cellular component or biogenesis (GO:0071840) with $n = 11$, the molecular function phosphatase activity (GO:0016791) with $n = 3$, and the cellular components for neuron part (GO:0097458) and cell projection (GO:0042995) each with $n = 3$. Though not significantly enriched—likely due to the especially large number of gene annotations contained within each GO term category—large numbers of *gas-1* RC line SNPs mapped to the following categories: biological processes of transport ($n = 9$) (GO:0006810), reproduction ($n = 7$) (GO:0000003), anatomical structure morphogenesis ($n = 5$) (GO:0009653), embryo development ($n = 8$) (GO:0009790), and metabolic process ($n = 19$) (GO:0008152); cellular components of membrane ($n = 14$) (GO:0016020), membrane part ($n = 12$) (GO:0044425), and intracellular organelle ($n = 7$) (GO:0043229); and the molecular functions of binding ($n = 19$) (GO:0005488), catalytic activity ($n = 16$) (GO:0003824), and ion binding ($n = 11$) (GO:0043167). Functional analysis of the *gas-1* interactome gene set (see Chapter 2) reveals shared enrichment of the three biological process categories that were enriched in the *gas-1* RC lines—cell organization and biogenesis, cytoskeleton organization and biogenesis, and protein modification—however, the molecular function category of phosphoprotein phosphatase activity, which was significantly enriched in the *gas-1* RC lines was not significantly enriched in the *gas-1* interactome gene set (Appendices C and D). A complete accounting of *gas-1* RC GO enrichment can be found in Appendix D.

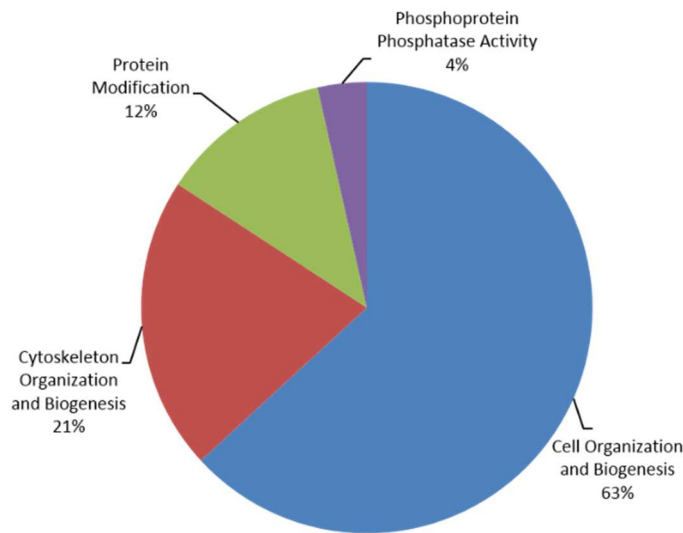


Figure 3.5. Proportions of GO slim categories significantly enriched among the *gas-1* RC. Proportions indicate share of GO terms contained within each GO slim category. See text for details on analysis and interpretation.

Discussion

In this chapter, I examined phenotypic and genetic patterns of fitness recovery of 24 *C. elegans gas-1 (fc21)* ETC-mutant lineages evolved in large population sizes; i.e., under conditions of efficient selection. My goal was to evaluate the capacity of new mutations to recover ancestral (wildtype) levels of fitness in a mitochondrial-deficient population. I showed that evolution in these favorable population genetic conditions can at least partly recover ancestral fitness and phenotypes through the accumulation of putatively beneficial and compensatory mutations across 60 generations, and that outcrossing drives faster fitness recovery.

Phenotypic analysis of *gas-1* G0 strain. In agreement with previous findings from our lab and others' (Kayser et al. 2004; Leu 2015), the *gas-1* G0 strain had significantly lower fitness than N2 in both absolute (W) and relative (ω) terms. I continue to observe longer

(though not significantly so here) lifespan in our backcrossed *gas-1* G0 strain compared to N2, in contrast to a previous study showing shorter lifespan for non-backcrossed *gas-1* (*fc21*) (Kayser et al. 2004). Combined with decrease steady-state ATP levels and low rates of pharyngeal pumping (a proxy for chronological age) relative to N2 (Leu 2015), my results of reduced fecundity, delayed reproduction, and longer life support *gas-1* as a long-lived mitochondrial mutant. In addition, *gas-1* G0's reproductive period is shifted a day later than wildtype N2. My lifespan findings for *gas-1* G0 in this study do not differ significantly from the previous *gas-1* MA study, though here *gas-1* G0 lives about a day less than in the MA study (data not shown; Leu 2016). Unlike the *gas-1* MA study previous study, however, our N2 lived longer in our RC fitness assay 14.56 days verses 12.97 days (Leu 2016; Chapter 2).

Fitness and life histories of gas-1 RC lines. I found that relatively short-term evolution (60 generations) in favorable population genetic conditions can facilitate at least partial fitness recovery of the *gas-1* (*fc21*) ETC-mutant. I also find that improvements in fitness and phenotype are potentially greater in the *gas-1* RC lines compared to the *gas-1* MA lines (Figure 2.1 vs. Figure 3.2A). Unlike, the *gas-1* MA lines (Chapter 2; Table 2.1), which experienced reduced variance for absolute and relative fitness compared to the *gas-1* G0 progenitor, the *gas-1* RC lines experiences increased variance for these traits compared to *gas-1* G0 (Table 3.1). This difference suggests that the evolutionary paths to fitness and phenotypic recovery under MA conditions are more narrowly defined than under efficient selection, presumably because efficient selection conditions unlock more molecular pathways to fitness increase.

Evolution rates and mutation characteristics. Evolutionary rate considers all evolutionary forces influencing inherited genome change across generations. We observed a rate of nuclear genome change approximately 3.4 times slower in the *gas-1* RC lines ($8.01 (\pm 0.75) \times 10^{-10}$) than observed in the *gas-1* MA lines [$2.12 \pm (0.34) \times 10^{-9}$ per nucleotide per generation] (Wernick 2016). The slower rate of evolutionary genomic change within the *gas-1* RC large population lineages suggests natural selection may be eliminating mutation-generated nuclear genome variation through purifying and/or directional selection.

Mutational relationships with life histories and phenotypes. There were no significant correlations between the total numbers of SNPs, exonic SNPs, intronic SNPs, intergenic SNPs, synonymous SNPs, non-synonymous SNPs, or *gas-1* interactor SNPs and fitness relative to *gas-1* G0 and fitness class—high or low, the observed presence of males, total reproductive output, or lifespan (data not shown).

Gene ontology analysis indicates more pathways to recovery in recovery lines than under MA. To further characterize the mutations that arose in the *gas-1* RC lines I performed a gene ontology (GO term) enrichment analysis. When contrasted with analyses of the *gas-1* MA lines (Chapter 2), these results offer three ideas about the functional bases of evolution in these two population genetic environments. First, though it is a small sample size of SNPs and total GO categories, fewer enriched GO slim categories emerged in *gas-1* MA lines than *gas-1* RC lines. This suggests that, under MA conditions, fewer evolutionary pathways are available that lead to fitness recovery. Second, there were no overlapping significantly enriched GO slim categories between the

two sets of lineages, suggesting that different evolutionary paths to fitness and phenotype recovery are available to lineages evolving under extreme drift versus efficient selection. In the *gas-1* RC lines, I saw that two biogenesis and organizational categories along with protein modification and phosphoprotein phosphatase were enriched. These categories are broader in their biological, cellular, and functional scope compared to the *gas-1* MA categories of transport and plasma membrane. Taken together, these patterns suggests that more pathways to recovery were available to *gas-1* RC as compared to *gas-1* MA lines. And third, we found evidence that several nuclear SNPs arising in the *gas-1* RC lines were annotated as being involved in fertility-related (and possibly embryogenesis) processes (Appendix D). This is in stark contrast to the *gas-1* MA line SNP set where we found no *gas-1* MA SNPs to be significantly enriched in embryonic development. It is possible that mutations in *gas-1* RC line fertility genes arose and were fixed simply owing to the large population sizes and consequently large target genome for beneficial mutation compared to *gas-1* MA – enabling positive selection.

Observations of males and implications for recovery. Here, I demonstrated that, among the six *gas-1* recovery lines with evidence for increased male frequency, experienced greater amounts of fitness recovery towards wildtype. It is important to note that our observations on the frequency of males in the lines are anecdotal and incomplete. More recovery lines may have generated males than we observed. The generation of males in *C. elegans* is a rare phenomenon (Riddle et al. 1997; Anderson et al. 2010). Outcrossing can play an important role in adaptive evolution by avoiding inbreeding depression and facilitating the spread of beneficial alleles (Anderson et al. 2010), especially in the

presence of high mutation loads (Morran et al. 2009) in accordance with evolutionary theory (Smith 1978). It is possible that our generational bleach age-sync treatments created an appropriate stressor for the generation of males; our laboratory strains are normally reared without such interference. This highlights the unexpected nature of experimental evolution in the lab.

Conclusion. I found that average fitness of evolved *gas-1* RC lines increased as expected, though not significantly as a group. Instead, two distinct groups of lines emerged—one that evolved significantly higher average fitness, and containing some lines that either recovered or exceeded wildtype fitness levels, and another group that showed more modest fitness non-significant gains. Interestingly, many lines in the first group were observed to generate significant numbers of males during experimental evolution, consistent with evolution of outcrossing either accompanying or driving rapid fitness recovery (Morran et al. 2009); these initial observations hold despite the finding of no increase non-disjunction rates and only two lines experiencing slightly elevated male frequency over wildtype. I then bioinformatically characterize the nuclear mutations that arose in the *gas-1* RC lines and find that potentially more paths to fitness recovery are available for large populations than small ones (see Chapter 2 GO Term characterization).

CHAPTER 4: CONCLUSION

The research presented here combines experimental evolution with phenotyping, whole-genome sequencing, and bioinformatic analyses to study the capacity of new mutations to yield adaptive evolution in *Caenorhabditis elegans*. Only a handful of studies have applied this combined approach in eukaryotic models to directly identify and test the effect of new mutations (Schlotterer et al. 2015). However, to our knowledge, this is one of only two studies (Joyner-Matos et al. 2011) that assesses the phenotypic consequences of evolution in the context of mitochondrial dysfunction. Together, the studies described in this thesis provide insight into the molecular genetic and phenotypic patterns of recovery of ETC-deficient populations—both large and small—over evolutionary time.

In Chapter 2, I demonstrate that the evolutionary fitness trajectory of a population clearly depends on its location on the fitness landscape (Wright 1932; Orr 2000; Poon and Otto 2000; Orr et al. 2003). Unlike traditional MA experiments with wildtype populations, fitness in the ETC-mutant *gas-1* MA lines improved with little increase in among-line variance combined with evidence for zygotic or post zygotic selection suggesting that selection—both purifying and positive—can be an extremely powerful force even in conditions of extreme genetic drift. Functional characterization of a putatively beneficial haplotype that arose in an MA experiment revealed successful identification of specific genetic changes can lead to a partial recovery of fitness and phenotype in an ETC-deficient mutant strain of *C. elegans*. My finding that the *gas-1*^{MA431} beneficial haplotype was associated with a decrease in steady-state endogenous ROS is consistent with decreased

ROS levels in *gas-1* MA in general (Lue 2015). Furthermore, my finding that steady-state ATP levels in the *gas-1*^{MA431} decreased compared to wildtype suggests that fermentation may be metabolic strategy to cope with increase oxidative stress (Lue 2015).

In Chapter 3, I found that the *gas-1* RC lines we had higher variance in fitness recovery compared to the *gas-1* MA lines suggesting that recovery experiments—with efficient selection—unlock more evolutionary paths to recovery. Specifically, in this experiment, the *gas-1* RC lines were capable of sexual recombination—a biological phenomenon not available to the *gas-1* MA lines. As a result, new combinations of mutation (and possibly initial genetic variation; Table 3.1) could be combined that aid in recovery. This is supported by the high fitness recovery in the *gas-1* RC lines observed to generate significant numbers of males during experimental evolution, consistent with evolution of outcrossing either accompanying or driving rapid fitness recovery (Morran et al. 2009). I then bioinformatically characterize the nuclear mutations that arose in the *gas-1* RC lines and find that potentially more paths to fitness recovery are available for large populations than small ones; and that positive selection may play a larger role in adaptive evolution in large populations than small ones.

Combined, the results of these two studies suggest that the *gas-1* (*fc21*) mutation is harmful to the point that few new deleterious mutations are viable. The majority of new mutations are deleterious for organismal fitness; therefore, the finding that new mutations led to an increase in unviable eggs in the *gas-1* MA lines poses a significant threat to survival. However, possibly owing to large clutch sizes, it appears the availability of beneficial or compensatory mutations visible to selection is high enough in *C. elegans*

that fitness can partially rebound even under extreme drift conditions. The magnitude of fitness recovery from *gas-1* (*fc21*) genetic insult is potentially even greater under efficient selection conditions when there is observational evidence for the presence of males in the lines. However, when the fitness results of all the *gas-1* RC lines are combined, we see no evidence for significantly increased fitness with respect to *gas-1* G0. The number of unhatched, nonviable eggs in *gas-1* RC was not significantly higher than *gas-1* G0. This suggests that there may be two different evolutionary mechanisms aiding in fitness recovery in the *gas-1* MA and *gas-1* RC lines, embryonic or gametic selection and sexual reproduction, respectively.

Our findings illuminate the functional and molecular genetic properties of mutations responsible for adaptation and compensation in ETC-deficient populations. Characterizing the phenotypic effects of the haplotype in the *gas-1*^{MA431} also contributes to our understanding of the epistatic mechanisms of compensatory adaptation in eukaryotes and how mitochondrial dysfunction can contribute to evolutionary dynamics. Future work will compare the phenotypic effects of these mutations with the phenotypic effects of candidate compensatory mutations that arose in the *gas-1* RC. From this information, we will be able to see different patterns in selective effects of mutations that compensate for ETC dysfunction under different population-genetic conditions. In the future, the results of these *gas-1* experiments may also be compared to the results of other strains in our ETC-mutant pipeline to see if there are similar patterns of fitness maintenance, compensatory evolution, and comparable patterns and rates of mutation. These findings are expected to be of broad interest and relevant to other eukaryotes since

few studies have addressed the population-level consequences of new mutation affecting mitochondrial ETC deficiencies in eukaryotic systems

REFERENCES

- Ameur, A., J. B. Stewart, C. Freyer, E. Hagström, M. Ingman, N. G. Larsson, and U. Gyllensten. 2011. Ultra-deep sequencing of mouse mitochondrial DNA: Mutational patterns and their origins. *PLoS Genet.* 7:18–21.
- Anderson, J. L., L. T. Morran, and P. C. Phillips. 2010. Outcrossing and the maintenance of males within *C. elegans* populations. *J. Hered.* 101:62–74.
- Baer, C. F., M. M. Miyamoto, and D. R. Denver. 2007. Mutation rate variation in multicellular eukaryotes: causes and consequences. *Nat. Rev. Genet.* 8:619–631.
- Baer, C. F., F. Shaw, C. Steding, M. Baumgartner, A. Hawkins, A. Houppert, N. Mason, M. Reed, K. Simonelic, W. Woodard, and M. Lynch. 2005. Comparative evolutionary genetics of spontaneous mutations affecting fitness in rhabditid nematodes. 102:5785–5790.
- Ballard, J. W. O., and M. C. Whitlock. 2004. The incomplete natural history of mitochondria. *Mol. Ecol.* 13:729–744.
- Blier, P. 2001. Natural selection and the evolution of mtDNA-encoded peptides: evidence for intergenomic co-adaptation. *Trends Genet.* 17:400–406.
- Brenner, S. 1974. The genetics of *Caenorhabditis elegans*. *Genetics* 77:71–94.
- Buckling, A., R. Craig Maclean, M. Brockhurst, and N. Colegrave. 2009. The Beagle in a bottle. *Nature* 457:824–9.
- Cipriani, P. G., and F. Piano. 2011. *Methods in Cell Biology*.
- Colangeli, R., V. L. Arcus, R. T. Cursons, A. Ruthe, N. Karalus, K. Coley, S. D. Manning, S. Kim, E. Marchiano, and D. Alland. 2014. Whole genome sequencing of *Mycobacterium tuberculosis* reveals slow growth and low mutation rates during latent infections in humans. *PLoS One* 9:1–9.
- Colegrave, N., and A. Buckling. 2005. Microbial experiments on adaptive landscapes. *Bioessays* 27:1167–73.
- Cooke, M. S., and M. D. Evans. 2007. 8-Oxo-deoxyguanosine: Reduce, reuse, recycle? *Proc Natl Acad Sci U S A* 104:13535–13536.
- Cooke, M. S., M. D. Evans, M. Dizdaroglu, and J. Lunec. 2003. Oxidative DNA damage: mechanisms, mutation, and disease. *FASEB J.* 17:1195–214.
- Crow, J. F., and M. Kimura. 1970. *An Introduction to Population Genetics Theory*. Harper and Row, New York, NY.
- Demple, B., and L. Harrison. 1994. Repair of oxidative damage to DNA: enzymology and biology. *Annu. Rev. Biochem.* 63:915–948.

- Denver, D. R., P. C. Dolan, L. J. Wilhelm, W. Sung, J. I. Lucas-Lledó, D. K. Howe, S. C. Lewis, K. Okamoto, W. K. Thomas, M. Lynch, and C. F. Baer. 2009. A genome-wide view of *Caenorhabditis elegans* base-substitution mutation processes. *Proc. Natl. Acad. Sci. U. S. A.* 106:16310–4.
- Denver, D. R., D. K. Howe, L. J. Wilhelm, C. a Palmer, J. L. Anderson, K. C. Stein, P. C. Phillips, and S. Estes. 2010. Selective sweeps and parallel mutation in the adaptive recovery from deleterious mutation in *Caenorhabditis elegans*. *Genome Res.* 20:1663–71.
- Dowling, D. K., U. Friberg, and J. Lindell. 2008. Evolutionary implications of non-neutral mitochondrial genetic variation. *Trends Ecol. Evol.* 23:546–54.
- Ellison, C. H. K. E. 2006. DISRUPTION OF MITOCHONDRIAL FUNCTION IN INTERPOPULATION HYBRIDS OF *TIGRIOPUS CALIFORNICUS*. 60:1382–1391.
- Equence, C. E. S., T. O. B. Iology, C. The, S. Consortium, and T. C. elegans S. Consortium. 1998. Genome Sequence of the Nematode *C. elegans*: A Platform for Investigating Biology. *Science* 282:2012–2018.
- Estes, S., A. L. Coleman-Hulbert, K. A. Hicks, G. de Haan, S. R. Martha, J. B. Knapp, S. W. Smith, K. C. Stein, and D. R. Denver. 2011a. Natural variation in life history and aging phenotypes is associated with mitochondrial DNA deletion frequency in *Caenorhabditis briggsae*. *BMC Evol. Biol.* 11:11.
- Estes, S., and M. Lynch. 2003. RAPID FITNESS RECOVERY IN MUTATIONALLY DEGRADED LINES OF. *Evolution (N. Y.)*. 57:1022–1030.
- Estes, S., P. C. Phillips, and D. R. Denver. 2011b. Fitness recovery and compensatory evolution in natural mutant lines of *C. elegans*. *Evolution (N. Y.)*. 65:2335–2344.
- Estes, S., P. C. Phillips, D. R. Denver, W. K. Thomas, and M. Lynch. n.d. Mutation Accumulation in Populations of Varying Size : The Distribution of Mutational Effects for Fitness Correlates in *Caenorhabditis elegans*. *Genetics* 1279:1269–1279.
- Eyre-Walker, a, and P. Keightley. 2007. The distribution of fitness effects of new mutations. *Nat. Rev. Genet.* 8:610–8.
- Fassone, E., and S. Rahman. 2012. Complex I deficiency: clinical features, biochemistry and molecular genetics. *J. Med. Genet.* 49:578–90.
- Fisher, R. 1930. *On the Genetical Theory of Natural Selection*. Oxford University Press, Oxford, UK.
- Fleury, C., B. Mignotte, and J.-L. Vayssière. 2002. Mitochondrial reactive oxygen species in cell death signaling. *Biochimie* 84:131–41.
- Ford, C. B., P. L. Lin, M. R. Chase, R. R. Shah, O. Iartchouk, J. Galagan, N. Mohaideen, T. R. Ioeberger, J. C. Sacchettini, M. Lipsitch, J. L. Flynn, and S. M. Fortune. 2011. Use of

- whole genome sequencing to estimate the mutation rate of *Mycobacterium tuberculosis* during latent infection. *Nat. Genet.* 43:482–6.
- Garcia, A. M., R. B. Calder, M. E. T. Doll, M. Lundell, P. Kapahi, and J. Vijg. 2010. Age- and temperature-dependent somatic mutation accumulation in *Drosophila melanogaster*. *PLoS Genet.* 6:27.
- Genova, M. L., B. Ventura, G. Giuliano, C. Bovina, G. Formiggini, G. Parenti Castelli, and G. Lenaz. 2001. The site of production of superoxide radical in mitochondrial Complex I is not a bound ubiquinone but presumably iron-sulfur cluster N2. *FEBS Lett.* 505:364–368.
- Gray, J. C., and A. D. Cutter. 2014. Mainstreaming *Caenorhabditis elegans* in experimental evolution. *Proc. Biol. Sci.* 281:20133055.
- Haag-Liautard, C., N. Coffey, D. Houle, M. Lynch, B. Charlesworth, and P. D. Keightley. 2008. Direct estimation of the mitochondrial DNA mutation rate in *Drosophila melanogaster*. *PLoS Biol.* 6:1706–1714.
- Halligan, D. L., and P. D. Keightley. 2009. Spontaneous Mutation Accumulation Studies in Evolutionary Genetics. *Annu. Rev. Ecol. Evol. Syst.* 40:151–172.
- Harcombe, W. R., R. Springman, and J. J. Bull. 2009. Compensatory evolution for a gene deletion is not limited to its immediate functional network. *BMC Evol. Biol.* 9:106.
- Hartl, D. L., and A. G. Clark. 2007. *Principles of Population Genetics*. Fourth. Sinauer Associates, Sunderland, MA, MA.
- Hartman, P. S., N. Ishii, E.-B. Kayser, P. G. Morgan, and M. M. Sedensky. 2001. Mitochondrial mutations differentially affect aging, mutability and anesthetic sensitivity in *Caenorhabditis elegans*. *Mech. Ageing Dev.* 122:1187–1201.
- Hirst, J., M. S. King, and K. R. Pryde. 2008. The production of reactive oxygen species by complex I. *Biochem. Soc. Trans.* 36:976–980.
- Holmström, K. M., and T. Finkel. 2014. Cellular mechanisms and physiological consequences of redox-dependent signalling. *Nat. Rev. Mol. Cell Biol.* 15:411–21. Nature Publishing Group.
- Hope, I. A. 1999. *C. elegans: A Practical Approach*. OUP Oxford.
- Hsie, A. W., L. Recio, D. S. Katz, C. Q. Lee, M. Wagner, and R. L. Schenley. 1986. Evidence for reactive oxygen species inducing mutations in mammalian cells. *Proc. Natl. Acad. Sci. U. S. A.* 83:9616–20.
- Hunte, C., V. Zickermann, and U. Brandt. 2010. Functional modules and structural basis of conformational coupling in mitochondrial complex I. *Science* 329:448–451.
- Jammes, F., C. Song, D. Shin, S. Munemasa, K. Takeda, D. Gu, D. Cho, S. Lee, R. Giordo, S.

- Sritubtim, N. Leonhardt, B. E. Ellis, Y. Murata, and J. M. Kwak. 2009. MAP kinases MPK9 and MPK12 are preferentially expressed in guard cells and positively regulate ROS-mediated ABA signaling. *Proc. Natl. Acad. Sci. U. S. A.* 106:20520–20525.
- Joseph, S. B., and D. W. Hall. 2004. Spontaneous mutations in diploid *Saccharomyces cerevisiae*: More beneficial than expected. *Genetics* 168:1817–1825.
- Joyner-Matos, J., L. C. Bean, H. L. Richardson, T. Sammeli, and C. F. Baer. 2011. No Evidence of Elevated Germline Mutation Accumulation Under Oxidative Stress in *Caenorhabditis elegans*. *Genetics* 189:1439–1447.
- Joyner-Matos, J., K. A. Hicks, D. Cousins, M. Keller, D. R. Denver, C. F. Baer, and S. Estes. 2013. Evolution of a Higher Intracellular Oxidizing Environment in *Caenorhabditis elegans* under Relaxed Selection. *PLoS One* 8:4–9.
- Kawecki, T. J., R. E. Lenski, D. Ebert, B. Hollis, I. Olivieri, and M. C. Whitlock. 2012. Experimental evolution. *Trends Ecol. Evol.* 27:547–60. Elsevier Ltd.
- Kayser, E.-B., P. G. Morgan, and M. M. Sedensky. 1999. Downloaded From: <http://anesthesiology.pubs.asahq.org/pdfaccess.ashx?url=/data/Journals/JASA/931258/> on 03/10/2016. *Anesthesiology* 1 90.
- Kayser, E.-B., M. M. Sedensky, and P. G. Morgan. 2004. The effects of complex I function and oxidative damage on lifespan and anesthetic sensitivity in *Caenorhabditis elegans*. *Mech. Ageing Dev.* 125:455–64.
- Kim, Y., K. Schumaker, and J.-K. Zhu. 2006. Ems Mutagenesis of *Arabidopsis*. *Methods Mol. Biol.* 323:101–103.
- Kimura, M. 1983. The Neutral Theory of Molecular Evolution.
- Kimura, M. 1985. The role of compensatory neutral mutations in molecular evolution. *J. Genet.* 64:7–19.
- Kondrashov, F. A., and A. S. Kondrashov. 2010. Measurements of spontaneous rates of mutations in the recent past and the near future. *Philos. Trans. R. Soc. B Biol. Sci.* 365:1169–1176.
- Lande, R. 1994. Risk of Population Extinction from Fixation of New Deleterious Mutations. *Evolution (N. Y.)* 48:1460–1469.
- Lanfear, R., H. Kokko, and A. Eyre-Walker. 2014. Population size and the rate of evolution. *Trends Ecol. Evol.* 29:33–41. Elsevier Ltd.
- Lenaz, G., R. Fato, M. L. Genova, C. Bergamini, C. Bianchi, and A. Biondi. 2006. Mitochondrial Complex I: structural and functional aspects. *Biochim. Biophys. Acta* 1757:1406–20.
- Lobo, I. 2008. Epistasis: Gene Interaction and the Phenotypic Expression of Complex

- Diseases Like Alzheimer's. *Nat. Educ.* 1:180.
- Lue, M. J. 2015. Phenotypic and Mutational Consequences of Mitochondrial ETC Genetic Damage. Portland State University.
- Lynch, M., J. Blanchard, D. Houle, T. Kibota, S. Schultz, L. Vassilieva, and Willis John. 1999. Perspective: Spontaneous deleterious mutations. *Evolution* (N. Y). 53:645–663.
- Lynch, M., I. J. Conery, and R. Burger. 1995a. Mutation Accumulation and the Extinction of Small Populations Author (s): Michael Lynch , John Conery and Reinhard Burger Source : *The American Naturalist* , Vol . 146 , No . 4 (Oct . , 1995), pp . 489-518 Published by : The University of Chicago Press . 146:489–518.
- Lynch, M., J. Conery, and R. Burger. 1995b. Mutation Accumulation and the Extinction of Small Populations. *Am. Nat.* 146:489–518.
- Lynch, M., W. Sung, K. Morris, N. Coffey, C. R. Landry, E. B. Dopman, W. J. Dickinson, K. Okamoto, S. Kulkarni, D. L. Hartl, and W. K. Thomas. 2008. A genome-wide view of the spectrum of spontaneous mutations in yeast. *Proc. Natl. Acad. Sci. U. S. A.* 105:9272–9277.
- MacLean, R. C., A. R. Hall, G. G. Perron, and A. Buckling. 2010. The population genetics of antibiotic resistance: integrating molecular mechanisms and treatment contexts. *Nat. Rev. Genet.* 11:405–14. Nature Publishing Group.
- McBride, R. C., C. B. Ogbunugafor, and P. E. Turner. 2008. Robustness promotes evolvability of thermotolerance in an RNA virus. *BMC Evol Biol* 8:231.
- Meer, M. V, A. S. Kondrashov, Y. Artzy-Randrup, and F. A. Kondrashov. 2010. Compensatory evolution in mitochondrial tRNAs navigates valleys of low fitness. *Nature* 464:279–82. Nature Publishing Group.
- Mittler, R., S. Vanderauwera, N. Suzuki, G. Miller, V. B. Tognetti, K. Vandepoele, M. Gollery, V. Shulaev, and F. Van Breusegem. 2011. ROS signaling: The new wave? *Trends Plant Sci.* 16:300–309. Elsevier Ltd.
- Morran, L. T., M. D. Parmenter, and P. C. Phillips. 2009. Mutation load and rapid adaptation favour outcrossing over self-fertilization. *Nature* 462:350–2.
- Murphy, M. P. 2009. How mitochondria produce reactive oxygen species. *Biochem. J.* 417:1–13.
- Nagao, T., S. Leuzinger, D. Acampora, a Simeone, R. Finkelstein, H. Reichert, and K. Furukubo-Tokunaga. 1998. Developmental rescue of *Drosophila* cephalic defects by the human *Otx* genes. *Proc. Natl. Acad. Sci. U. S. A.* 95:3737–42.
- Nijhuis, M., R. Schuurman, D. de Jong, J. Erickson, E. Gustchina, J. Albert, P. Schipper, S. Gulnik, and C. Boucher. 1999. Increased fitness of drug resistant HIV-1 protease as

- a result of acquisition of compensatory mutations during suboptimal therapy. *AIDS* 13:2349–2359.
- Ohta, T. 1992. The nearly neutral theory of molecular evolution. *Annu. Rev. Ecol. Evol. Syst.* 23:263–286.
- Orr, H. A. 2000. Adaptation and the cost of complexity. 54:13–20.
- Orr, H. A., K. C. Atwood, L. K. Schnieder, F. J. Ryan, J. Bull, M. Badgett, H. Wichman, C. Chatfield, C. M. Theobald, L. de Haan, P. Gerrish, P. J. Gerrish, R. E. Lenski, J. H. Gillespie, J. H. Gillespie, J. H. Gillespie, E. J. Gumbel, J. B. S. Haldane, J. B. S. Haldane, D. L. Hartl, C. H. Taubes, K. Holder, J. Bull, M. Imhof, C. Schlotterer, M. Kimura, M. R. Leadbetter, G. Lindgren, H. Rootzen, R. E. Lenski, M. Travisano, J. M. Smith, H. A. Orr, H. A. Orr, H. A. Orr, H. A. Orr, S. P. Otto, C. D. Jones, D. E. Rozen, J. A. G. M. de Visser, P. J. Gerrish, L. M. Wahl, D. C. Krakauer, I. Weissman, H. A. Wichman, M. R. Badgett, L. A. Scott, C. M. Boulianne, and J. J. Bull. 2003. The distribution of fitness effects among beneficial mutations. *Genetics* 163:1519–26.
- Osada, N., and H. Akashi. 2011. 1 . Division of Evolutionary Genetics , National Institute of Genetics , 1111 Yata , Department of Genetics , The Graduate University for Advanced Studies (SOKENDAI), 1111 , Yata , Mishima , Shizuoka 411-8540 , Japan .
* Corresponding author : Division .
- Peters, A. D., D. L. Halligan, M. C. Whitlock, and P. D. Keightley. 2003. Dominance and Overdominance of Mildly Deleterious Induced Mutations for Fitness Traits in *Caenorhabditis elegans*. *Genetics* 165:589–599.
- Phillips, W. S., E. S. Coleman-Hulbert, Anna L. Weiss, D. K. Howe, S. Ping, R. I. Wernick, S. Estes, and D. R. Denver. 2014. Population size mediates evolutionary conflict between nematodes and their mitochondrial DNA. *PLoS Biol.* Submiss.
- Poelwijk, F. J., D. J. Kiviet, D. M. Weinreich, and S. J. Tans. 2007. Empirical fitness landscapes reveal accessible evolutionary paths. *Nature* 445:383–6.
- Polyak, K., Y. Li, H. Zhu, C. Lengauer, J. K. Willson, S. D. Markowitz, M. a Trush, K. W. Kinzler, and B. Vogelstein. 1998. Somatic mutations of the mitochondrial genome in human colorectal tumours. *Nat. Genet.* 20:291–293.
- Poon, A., and L. Chao. 2005. The rate of compensatory mutation in the DNA bacteriophage phiX174. *Genetics* 170:989–99.
- Poon, A., and S. P. Otto. 2000. COMPENSATING FOR OUR LOAD OF MUTATIONS: FREEZING THE MELTDOWN OF SMALL POPULATIONS. *Evolution* (N. Y). 54:1467.
- Riddle, D. L., B. Thomas, B. J. Meyer, and J. R. Priess. 1997. *C. elegans* II.
- Rokyta, D., M. R. Badgett, I. J. Molineux, and J. J. Bull. 2002. Experimental Genomic Evolution : Extensive Compensation for Loss of DNA Ligase Activity in a Virus. *Mol.*

- Biol. Evol. 230–238.
- Rutter, M. T., F. H. Shaw, and C. B. Fenster. 2010. Spontaneous mutation parameters for *Arabidopsis thaliana* measured in the wild. *Evolution* 64:1825–35.
- Saxer, G., P. Havlak, S. A. Fox, M. A. Quance, S. Gupta, Y. Fofanov, J. E. Strassmann, and D. C. Queller. 2012. Whole Genome Sequencing of Mutation Accumulation Lines Reveals a Low Mutation Rate in the Social Amoeba *Dictyostelium discoideum*. *PLoS One* 7:e46759.
- Sazanov, L. A. 2015. REVIEWS A giant molecular proton pump : structure and mechanism of respiratory complex I. *Nat. Publ. Gr.* 16:375–388. Nature Publishing Group.
- Schlotterer, C., R. Kofler, E. Versace, R. Tobler, and S. U. Franssen. 2015. Combining experimental evolution with next-generation sequencing: a powerful tool to study adaptation from standing genetic variation. *Heredity (Edinb)*. 114:431–440.
- Shaw, F. H., C. J. Geyer, and R. G. Shaw. 2002. A comprehensive model of mutations affecting fitness and inferences for *Arabidopsis thaliana*. *Evolution* 56:453–63.
- Stiernagle, T. 2006. Maintenance of *C. elegans*. *WormBook*, Minneapolis, Minnesota.
- Taylor, D. R., C. Zeyl, and E. Cooke. 2002. Conflicting levels of selection in the accumulation of mitochondrial defects in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U. S. A.* 99:3690–4.
- Taylor, R. W., and D. M. Turnbull. 2005. Mitochondrial DNA mutations in human disease. *Nat Rev Genet* 6:389–402.
- Vanderauwera, S., N. Suzuki, G. Miller, B. van de Cotte, S. Morsa, J.-L. Ravanat, A. Hegie, C. Triantaphylidès, V. Shulaev, M. C. E. Van Montagu, F. Van Breusegem, and R. Mittler. 2011. Extranuclear protection of chromosomal DNA from oxidative stress. *Proc. Natl. Acad. Sci. U. S. A.* 108:1711–1716.
- Vassilieva, L. L., A. M. Hook, M. Lynch, and S. Url. 2000. The Fitness Effects of Spontaneous Mutations in *Caenorhabditis elegans*. *Society* 54:1234–1246.
- Vinothkumar, K. R., J. Zhu, and J. Hirst. 2014. Architecture of mammalian respiratory complex I. *Nature* 515:80–84. Nature Publishing Group.
- Wernick, R. I., S. Estes, D. K. Howe, and D. R. Denver. 2016. Paths of Heritable Mitochondrial DNA Mutation and Heteroplasmy in Reference and *gas-1* Strains of *Caenorhabditis elegans*. *Front. Genet.* 7:1–12.
- Wikström, M. 1984. Two protons are pumped from the mitochondrial matrix per electron transferred between NADH and ubiquinone. *FEBS Lett.* 169:300–304.
- Wirth, C., U. Brandt, C. Hunte, V. Zickermann, and V. Zick. 2016. Structure and function

- of mitochondrial complex I. *BBA - Bioenerg.* 1857:902–914. Elsevier B.V.
- Wiseman, H., and B. Halliwell. 1996. Damage to DNA by reactive oxygen and nitrogen species: role in inflammatory disease and progression to cancer. *Biochem. J.* 313 (Pt 1:17–29.
- Wood, W. B. 1988. *The Nematode Caenorhabditis elegans*. Cold Spring Harbor Monograph Archive.
- Wright, S. 1968. *Evolution and the Genetics of Populations*, Vol. 1.
- Wright, S. 1932. The roles of mutation, inbreeding, crossbreeding and selection in evolution.
- Zeeberg, B. R., W. Feng, G. Wang, M. D. Wang, A. T. Fojo, M. Sunshine, S. Narasimhan, D. W. Kane, W. C. Reinhold, S. Lababidi, K. J. Bussey, J. Riss, J. C. Barrett, and J. N. Weinstein. 2003. GoMiner: a resource for biological interpretation of genomic and proteomic data. *Genome Biol.* 4:R28.
- Zeeberg, B. R., H. Qin, S. Narasimhan, M. Sunshine, H. Cao, D. W. Kane, M. Reimers, R. M. Stephens, D. Bryant, S. K. Burt, E. Elnekave, D. M. Hari, T. a Wynn, C. Cunningham-Rundles, D. M. Stewart, D. Nelson, and J. N. Weinstein. 2005. High-Throughput GoMiner, an “industrial-strength” integrative gene ontology tool for interpretation of multiple-microarray experiments, with application to studies of Common Variable Immune Deficiency (CVID). *BMC Bioinformatics* 6:168.
- Zhang, L., and L. T. Watson. 2008. Analysis of the fitness effect of compensatory mutations. *HFSP J.* 3:47–54.
- Zhong, W., and P. W. Sternberg. 2006. *Genome-Wide Prediction*. 380.
- Zickermann, V., C. Wirth, H. Nasiri, K. Siegmund, H. Schwalbe, C. Hunte, and U. Brandt. 2015. Mechanistic insight from the crystal structure of mitochondrial complex I. *Mech. insight from Cryst. Struct. mitochondrial complex I* 347:44–49.

APPENDIX A: GO term data for *gas-1* MA lines

GO slim enrichment categories for <i>gas-1</i> MA – biological process				
Go Slim Category	GO Term Count	Gene Count	Genes	
Transport	29	7	<i>daf-2, sel-2, smf-3, twk-31, F41E7.2, rheb-1, F23H11.2</i>	
Protein Transport	14	2	<i>daf-2, rheb-1</i>	
Metabolism	14	8	<i>F23H11.2, C04E12.10, daf-2, mlt-11, T05A7.11, F37H8.3, ZC8.6, daf-2</i>	
Cell Communication	12	4	<i>W02B8.2, sel-2, rheb-1, daf-2</i>	
Cell Organization and Biogenesis	10	4	<i>sel-2, tbcc-1, F23H11.2, daf-2</i>	
Biosynthesis	9	2	<i>T05A7.11, F23H11.2</i>	
Signal Transduction	9	4	<i>W02B8.2, sel-2, rheb-1, daf-2</i>	
Nucleobase, Nucleoside, Nucleotide and Nucleic Acid Metabolism	9	2	<i>T05A7.11, F23H11.2</i>	
Response to Endogenous Stimulus	8	1	<i>daf-2</i>	
Response to Stress	6	2	<i>smf-3, daf-2</i>	
Response to External Stimulus	5	1	<i>daf-2</i>	
Organelle Organization and Biogenesis	5	1	<i>daf-2</i>	
Catabolism	4	2	<i>C04E12.10, daf-2</i>	
Ion Transport	2	3	<i>smf-3, twk-31, F41E7.2</i>	
Protein Metabolism	2	4	<i>mlt-11, C04E12.10, T05A7.11, daf-2</i>	
Development	2	6	<i>W02B8.2, sel-2, mlt-11, twk-31, ZC8.6, daf-2</i>	

GO Slim Categories in bold indicate $\alpha = 0.05$ significance for functional enrichment in the GoMiner output (see Methods). Genes in bold indicate their presence in any significantly enriched GO child category at $\alpha = 0.05$ for that particular GO Slim Category (p values not shown).

GO slim enrichment categories for *gas-1* MA – cellular component

GO Slim Category	GO Term Count	Gene Count	Genes
Cell	2	8	<i>rheb-1</i> , <i>twk-31</i> , <i>daf-2</i> , <i>smf-3</i> , <i>W02B8.2</i> , <i>C04E12.10</i> , <i>F23H11.2</i> , <i>sel-2</i>
Plasma Membrane	1	4	<i>rheb-1</i> , <i>twk-31</i> , <i>daf-2</i> , <i>smf-3</i>

GO Slim Categories in bold indicate $\alpha = 0.05$ significance for functional enrichment in the GoMiner output (see Methods). Genes in bold indicate their presence in any significantly enriched GO child category at $\alpha = 0.05$ for that particular GO Slim Category (p values not shown).

GO Slim enrichment categories for *gas-1* MA – molecular function

GO Slim Category	GO Term Count	Gene Count	Genes
Binding	9	5	<i>W02B8.2</i> , <i>rheb-1</i> , <i>tbcc-1</i> , <i>daf-2</i> , <i>F23H11.2</i>
Transporter Activity	7	3	<i>F41E7.2</i> , <i>smf-3</i> , <i>twk-31</i>
Catalytic Activity	6	5	<i>T05A7.11</i> , <i>F37H8.3</i> , <i>daf-2</i> , <i>ZC8.6</i> , <i>alh-2</i>
DNA Binding	5	1	<i>F23E7.2</i>
Nucleic Acid Binding	5	1	<i>F23E7.2</i>
Kinase Activity	3	1	<i>daf-2</i>
Receptor Activity	3	1	<i>daf-2</i>
Protein Kinase Activity	3	1	<i>daf-2</i>
Transferase Activity	3	1	<i>daf-2</i>
Signal Transducer Activity	3	1	<i>daf-2</i>
Transcription Factor Activity	3	1	<i>F23E7.2</i>
Protein Binding	2	2	<i>tbcc-1</i> , <i>daf-2</i>

GO Slim Categories in bold indicate $\alpha = 0.05$ significance for functional enrichment in the GoMiner output (see Methods). Genes in bold indicate their presence in any significantly enriched GO child category at $\alpha = 0.05$ for that particular GO Slim Category (p values not shown).

APPENDIX B: GO term data for N2 MA lines

GO slim enrichment categories for Baer N2 MA – biological process			
GO Slim Category	GO Term	Gene Count	Genes
Metabolism	31	23	Y39G10AR.11, hum-6, nek-4, <i>rrf-3</i> , <i>lagr-1</i> , T09B4.8, <i>smi-1</i> , <i>ptp-1</i> , Y45G12B.3, <i>fmo-3</i> , <i>cgt-2</i> , R02D3.3, <i>mix-1</i> , <i>rmf-113</i> , F11A3.2 , <i>mpk-1</i> , F11A3.2 , <i>rrf-3</i> , <i>lagr-1</i> , <i>let-526</i> , <i>rab-10</i> , <i>cgt-2</i> , R02D3.3
Biosynthesis	15	7	F11A3.2 , <i>rrf-3</i> , <i>lagr-1</i> , <i>let-526</i> , <i>rab-10</i> , <i>cgt-2</i> , R02D3.3
Cell Organization and Biogenesis	14	12	<i>rmf-113</i> , <i>bbs-4</i> , <i>rrf-3</i> , <i>smi-1</i> , <i>let-526</i> , <i>coh-4</i> , <i>rab-10</i> , F22E5.12 , <i>srz-80</i> , <i>rrf-3</i> , <i>pitp-1</i> , <i>fmo-3</i> , <i>hrp-2</i> , <i>rsa-2</i> , <i>vab-2</i> , R02D3.3 , <i>rmf-113</i> , F11A3.2 , <i>mpk-1</i> , <i>let-526</i> , <i>rab-10</i> , <i>coh-4</i> , <i>srz-80</i>
Reproduction	13	14	<i>rrf-3</i> , <i>pitp-1</i> , <i>fmo-3</i> , <i>hrp-2</i> , <i>rsa-2</i> , <i>vab-2</i> , R02D3.3 , <i>rmf-113</i> , F11A3.2 , <i>mpk-1</i> , <i>let-526</i> , <i>rab-10</i> , <i>coh-4</i> , <i>srz-80</i>
Organelle Organization and	13	8	<i>rmf-113</i> , <i>bbs-4</i> , <i>rrf-3</i> , <i>let-526</i> , <i>coh-4</i> , <i>srz-80</i> , <i>rsa-2</i> , <i>mix-1</i>
Cell Cycle	13	7	<i>rmf-113</i> , <i>mpk-1</i> , <i>rrf-3</i> , <i>coh-4</i> , <i>srz-80</i> , <i>rsa-2</i> , <i>mix-1</i>
Lipid Metabolism	12	2	<i>lagr-1</i> , <i>cgt-2</i>
Carbohydrate Metabolism	5	1	<i>aagr-3</i>
Catabolism	5	4	<i>mok-1</i> , <i>aagr-3</i> , <i>cpr-8</i> , F15D4.4
Development	4	17	Y39G10AR.11 , <i>bbs-4</i> , <i>rrf-3</i> , <i>smi-1</i> , <i>hrp-2</i> , <i>rsa-2</i> , <i>vab-2</i> , <i>mix-1</i> , R02D3.3 , T28D6.4 , <i>rmf-113</i> , F11A3.2 , <i>mpk-1</i> , <i>aagr-3</i> , <i>let-526</i> , <i>rab-10</i> , <i>srz-80</i>

Regulation of Gene Expression, Epigenetic	3	2	<i>rrf-3, mix-1</i>
Nucleobase, Nucleoside, Nucleotide	2	4	<i>rrf-3, smi-1, let-526, R02D3.3</i>
Embryonic Development	2	16	<i>Y39G10AR.11, bbs-4, rrf-3, smi-1, hrp-2, rsa-2, vab-2, R02D3.3, mix-1, rnf-113, F11A3.2, mpk-1, aagr-3, let-526, rab-10, srz-80</i>
Transport	1	10	<i>F11A3.2, bbs-4, best-8, dlt-1, let-526, rab-10, srz-80, eal-23, R02D3.3.</i>
Morphogenesis	1	4	<i>R02D3.3, vab-2, srz-80, let-526</i>
Signal Transduction	1	6	<i>B0334.6, rab-10, vab-2, mpk-1, hum-6, D1014.2</i>
Cell Communication	1	6	<i>B0334.6, rab-10, vab-2, mpk-1, hum-6, D1014.2</i>
Generation of Precursor Metabolites	1	1	<i>aagr-3</i>
Protein Metabolism	1	8	<i>rnf-113, nekl-4, F11A3.2, mpk-1, ptp-1, rab-10, cpr-8, F15D4.4</i>
Protein Biosynthesis	1	2	<i>F11A3.2, rab-10</i>

GO Slim Categories in bold indicate $\alpha = 0.05$ significance for functional enrichment in the GoMiner output (see Methods). Genes in bold indicate their presence in any significantly enriched GO child category at $\alpha = 0.05$ for that particular GO Slim Category (p values not shown).

GO slim enrichment categories for Baer N2 MA – cellular component

GO Slim Category	GO Term Count	Gene Count	Genes
Cell	41	27	<i>hrp-2, hum-6, ptp-1, smi-1, rsa-2, pitp-1, cpr-8, coh-4, lagr-1, F15D4.4, rnf-113, Y45G12B.3, F11A3.2, unc-80, best-8, R02D3.3, vab-2, mpk-1, aagr-3, nekl-4, T09B4.8, egl-23, rab-10, fmo-3, bbs-4, let-526, mix-1</i>
Intracellular	29	23	<i>hrp-2, hum-6, R02D3.3, ptp-1, smi-1, rsa-2, pitp-1, cpr-8, coh-4, mpk-1, aagr-3, nekl-4, T09B4.8, rab-10, lagr-1, F15D4.4, rnf-113, Y45G12B.3, fmo-3, F11A3.2, bbs-4, let-526, mix-1</i>
Cytoplasm	11	15	<i>hum-6, ptp-1, smi-1, cpr-8, mpk-1, aagr-3, nekl-4, T09B4.8, rab-10, lagr-1, F15D4.4, Y45G12B.3, fmo-3, F11A3.2, bbs-4</i>
Cytoskeleton	8	5	<i>hum-6, ptp-1, rsa-2, bbs-4, nekl-4</i>
Plasma Membrane	5	4	<i>best-8, vab-2, bbs-4, egl-23</i>
Nucleus	4	9	<i>hrp-2, R02D3.3, smi-1, rnf-113, mpk-1, coh-4, let-526, nekl-4, mix-1</i>
Microtubule Organizing Center	3	3	<i>rsa-2, bbs-4, nekl-4</i>
Cytoplasmic Membrane-Bound	3	1	<i>rab-10</i>
Chromosome	3	2	<i>coh-4, mix-1</i>
Vacuole	3	4	<i>rab-10, F15D4.4, cpr-8, aagr-3</i>
Cilium	2	2	<i>pitp-1, bbs-4</i>
Lysosome	1	3	<i>F15D4.4, cpr-8</i>
Golgi Apparatus	1	1	<i>rab-10</i>

GO Slim Categories in bold indicate $\alpha = 0.05$ significance for functional enrichment in the GoMiner output (see Methods). Genes in bold indicate their presence in any significantly enriched GO child category at $\alpha = 0.05$ for that particular GO Slim Category (p values not shown).

GO slim enrichment categories for Baer N2 MA – molecular function

GO Slim Category	GO Term Count	Gene Count	Genes
Catalytic Activity	20	19	<i>cpr-8, hum-6, lagr-1, bbs-4, aagr-3, Y45G12B.3, mpk-1, ptp-1, dhc-4, cgt-2, oac-20, rrf-3, fmo-3, T09B4.8, nekl-4, rnf-113, F15D4.4, F11A3.2, rab-10</i>
Hydrolase Activity	7	8	<i>hum-6, bbs-4, dhc-4, F15D4.4, cpr-8, aagr-3, ptp-1, rab-10</i>
Transferase Activity	7	8	<i>rrf-3, mpk-1, cgt-2, T09B4.8, lagr-1, oac-20, nekl-4, rnf-113</i>
Binding	3	16	<i>rsa-2, vab-2, mix-1, pitp-1, hum-6, aagr-3, mpk-1, ptp-1, fmo-3, smi-1, T09B4.8, nekl-4, hrp-2, rnf-113, let-526, rab-10</i>
Peptidase Activity	2	2	<i>F15D4.4, cpr-8</i>
Motor Activity	2	3	<i>hum-6, bbs-4, dhc-4</i>
Protein Binding	2	9	<i>rsa-2, vab-2, mix-1, hum-6, smi-1, T09B4.8, mpk-1, ptp-1, rab-10</i>
Kinase Activity	1	2	<i>mpk-1, nekl-4</i>
Receptor Binding	1	1	<i>vab-2</i>
Protein Kinase Activity	1	2	<i>mpk-1, nekl-4</i>
Signal Transducer Activity	1	3	<i>mpk-1, B0334.6, D1014.2</i>

GO Slim Categories in bold indicate $\alpha = 0.05$ significance for functional enrichment in the GoMiner output (see Methods). Genes in bold indicate their presence in any significantly enriched GO child category at $\alpha = 0.05$ for that particular GO Slim Category (p values not shown).

APPENDIX C: GO term data for *gas-1* centric interactome

GO slim enrichment categories for *gas-1* centric interactome – biological process

GO Slim Category	GO Term Count	% GO Terms	Gene Count
Metabolism	840	22.51%	2735
Cell organization and biogenesis	315	8.44%	1004
Nucleobase, nucleoside, nucleotide and nucleic acid metabolism	289	7.75%	993
Biosynthesis	241	6.46%	1069
Development	163	4.37%	2194
Protein metabolism	158	4.23%	981
Transport	155	4.15%	1012
Organelle organization and biogenesis	152	4.07%	687
Catabolism	123	3.30%	429
Cell differentiation	114	3.06%	619
Cell communication	108	2.89%	562
Cell cycle	106	2.84%	390
Response to stress	93	2.49%	501
Protein modification	91	2.44%	556
Morphogenesis	81	2.17%	653
Signal transduction	77	2.06%	489
Reproduction	74	1.98%	1514
Lipid metabolism	55	1.47%	268
Cytoskeleton organization and biogenesis	53	1.42%	225
Carbohydrate metabolism	44	1.18%	156
Ion transport	40	1.07%	271
Response to external stimulus	35	0.94%	213
Growth	34	0.91%	342
DNA metabolism	32	0.86%	164
Cell death	32	0.86%	397
Embryonic development	28	0.75%	1552
Behavior	24	0.64%	335
Cell-cell signaling	23	0.62%	86
Cell growth	21	0.56%	62
Protein transport	20	0.54%	182
Response to abiotic stimulus	19	0.51%	118
Generation of precursor metabolites and energy	19	0.51%	80
Cell homeostasis	17	0.46%	87
Regulation of gene expression, epigenetic	11	0.29%	101
Response to endogenous stimulus	10	0.27%	60
Protein biosynthesis	10	0.27%	204
Cell proliferation	9	0.24%	110
Mitochondrion organization and biogenesis	8	0.21%	63
Cell recognition	3	0.08%	44
Response to biotic stimulus	3	0.08%	52
Cytoplasm organization and biogenesis	1	0.03%	8

GO Slim Categories in bold indicate $\alpha = 0.05$ significance for functional enrichment in the GoMiner output (see Methods).

GO slim enrichment categories for *gas-1* centric interactome – cellular component

Category	GO Term Count	% GO Terms	SNP Count
Cell	326	31.38%	2612
Intracellular	290	27.91%	2412
Cytoplasm	130	12.51%	1539
Nucleus	78	7.51%	948
Cytoplasmic membrane-bound vesicle	25	2.41%	64
Nucleoplasm	25	2.41%	135
Chromosome	24	2.31%	136
Cytoskeleton	21	2.02%	161
Vacuole	18	1.73%	100
Mitochondrion	16	1.54%	268
Plasma membrane	15	1.44%	322
Nuclear chromosome	12	1.15%	69
Golgi apparatus	12	1.15%	104
Endoplasmic reticulum	9	0.87%	180
Endosome	8	0.77%	50
Ribosome	7	0.67%	85
Cytosol	7	0.67%	212
Nucleolus	4	0.38%	50
Nuclear membrane	2	0.19%	39
Extracellular region	2	0.19%	88
Microtubule organizing center	2	0.19%	39
Lysosome	2	0.19%	36
Extracellular matrix (sensu Metazoa)	2	0.19%	25
Peroxisome	1	0.10%	32
Lipid particle	1	0.10%	11

GO Slim Categories in bold indicate $\alpha = 0.05$ significance for functional enrichment in the GoMiner output (see Methods).

GO slim enrichment categories for *gas-1* centric interactome – molecular function

Category	GO Term Count	% GO Terms	SNP Count
Catalytic activity	201	31.90%	1967
Binding	120	19.05%	2306
Hydrolase activity	60	9.52%	695
Transferase activity	58	9.21%	676
Protein binding	39	6.19%	932
Transporter activity	26	4.13%	30
Nucleic acid binding	21	3.33%	664
Kinase activity	17	2.70%	253
Nucleotide binding	15	2.38%	747
DNA binding	11	1.75%	357
Protein kinase activity	9	1.43%	172
RNA binding	8	1.27%	231
Peptidase activity	5	0.79%	129
Cytoskeletal protein binding	5	0.79%	85
Enzyme regulator activity	5	0.79%	85
Nuclease activity	5	0.79%	55
Lipid binding	5	0.79%	51
Antioxidant activity	4	0.63%	29
Signal transducer activity	3	0.48%	119
Motor activity	2	0.32%	31
Transcription factor activity	2	0.32%	149
Translation factor activity, nucleic acid binding	2	0.32%	38
Carbohydrate binding	2	0.32%	26
Structural molecule activity	2	0.32%	145
Receptor binding	1	0.16%	72
Chromatin binding	1	0.16%	30
Actin binding	1	0.16%	47

GO Slim Categories in bold indicate $\alpha = 0.05$ significance for functional enrichment in the GoMiner output (see Methods).

APPENDIX D: GO term data for *gas-1* RC lines

GO slim enrichment categories for Baer N2 MA – biological process

GO Slim Category	GO Term Count	Gene Count	Gene
Cell organization and biogenesis	36	10	<i>TTN-1, SYD-2, SLT-1, TIR-1, FER-1, PPFR-1, MATH-33, CEN-12, Y34R4A, PPH-4.1</i>
Transport	23	9	<i>SYD-2, Y82E9BR.2, CED-12, Y34B4A.2, TWK-17, Y41E3.7, TIR-1, MRP-7, LGC-29</i>
Metabolism	23	19	<i>F10F2.2, Y54F10BM.3, PPFR-1, Y106G6D.4, MATH-33, PPH-4.1, SET-29, TTN-1, TIR-1, C17H12.3, Y82E9BR.2, ATIC-1, Y16B4A.2, ABHD-11.2, MRP-7, K07C5, NAC-43, Y34R4A, FKH-9</i>
Organelle organization and biogenesis	15	5	<i>TTN-1, PPFR-1, CED-12, Y34B4A.2, PPH-4.1</i>
Cell differentiation	12	5	<i>TTN-1, SLT-1, TIR-1, FER-1, FKH-9</i>
Cytoskeleton organization and biogenesis	12	4	<i>TTN-1, PPFR-1, CED-12, PPH-4.1</i>
Development	10	11	<i>SLT-1, PPFR-1, PPH-4.1, TTN-1, TIR-1, NLP-20, FER-1, T19D12.1, MATH-33, CED-12, F26F2.7</i>
Cell cycle	10	2	<i>PPFR-1, PPH-4.1</i>
Protein metabolism	9	10	<i>TTN-1, TIR-1, C17H12.3, Y54F10BM.3, PPFR-1, Y106G6D.4, MATH-33, Y16B4A.2, PPH-4.1, SET-29</i>
Morphogenesis	8	5	<i>TTN-1, SLT-1, TIR-1, CED-12, FKH-9</i>
Biosynthesis	7	4	<i>TIR-1, F10F2.2, ATIC-1, FKH-9</i>
Response to external stimulus	7	2	<i>TIR-1, SLT-1</i>

Nucleobase, nucleoside, nucleotide and nucleic acid metabolism	7	5	<i>TIR-1, F10F2.2, ATIC-1, Y82E9BR.2, FKH-9</i>
Protein modification			
	7	9	<i>TTN-1, TIR-1, C17H12.3, Y54F10BM.3, PPFR-1, Y106G6D.4, MATH-33, PPH-4.1, SET-29</i>
Reproduction	6	7	<i>PPFR-1, SYD-2, TIR-1, FER-1, CED-12, ATIC-1, PPH-4.1</i>
Cell communication	5	5	<i>SYD-2, TIR-1, AC3.12, SLT-1, NLP-20</i>
Growth	5	3	<i>SLT-1, TIR-1, MATH-33</i>
Cell growth	4	1	<i>SLT-1</i>
Signal transduction	3	4	<i>SLT-1, TIR-1, AC3.12, NLP-20</i>
Response to stress	1	1	<i>TIR-1</i>
Response to biotic stimulus	1	1	<i>TIR-1</i>
Cell-cell signaling	1	2	<i>TIR-1, SYD-2</i>

GO Slim Categories in bold indicate $\alpha = 0.05$ significance for functional enrichment in the GoMiner output (see Methods). Genes in bold indicate their presence in any significantly enriched GO child category at $\alpha = 0.05$ for that particular GO Slim Category (p values not shown).

GO slim enrichment categories for Baer N2 MA – cellular component

GO Slim Category	GO Term Count	Gene Count	Genes
Cell	17	14	SYD-2, FER-1, CED-12, TTN-1, TIR-1, Y82E9BR.2, MATH-33, PPFR-1, PPH-4.1, F10F2.2, SET-29, Y34B4A.2, FKH-9, TWK-17
Intracellular	9	12	TTN-1, CED-12, Y82E9BR.2, TIR-1, PPFR-1, SYD-2, MATH-33, PPH-4.1,
Cytoplasm	4	9	TTN-1, TIR-1, PPFR-1, SYD-2, MATH-33, PPH-4.1, F10F2.2, Y34B4A.2, CED-12
Cytoskeleton	3	2	TTN-1, PPH-4.1
Nucleus	3	4	Y82E9BR.2, MATH-33, SET-29, FKH-9
Plasma membrane	3	3	FER-1, TWK-17, CED-12

GO Slim Categories in bold indicate $\alpha = 0.05$ significance for functional enrichment in the GoMiner output (see Methods). Genes in bold indicate their presence in any significantly enriched GO child category at $\alpha = 0.05$ for that particular GO Slim Category (p values not shown).

GO slim enrichment categories for Baer N2 MA – molecular function

GO Slim Category	GO Term Count	Gene Count	Genes
Binding	22	19	CED-12, PPFR-1, SYD-2, FER-1, Y41E3.7, TTN-1, TIR-1, SLT-1, MATH-33, TAG-80, Y34B4A.2, F10F2.2, Y106G6D.4, MRP-7, PPH-4.1, Y82E9BR.2, C52E12.1, FKH-9, F16A11.1
Protein binding	14	10	TIR-1, CED-12, PPFR-1, SLT-1, MATH-33, SYD-2, TTN-1, TAG-80, PPH-4.1, Y34B4A.2
Catalytic activity	12	16	F10F2.2, ATIC-1, PPH-4.1, Y82E9BR.2, C17H12.3, Y54F10BM.3, AC3.12, MATH-33, MRP-7, ABHD-11.2, K07C5.2, TTN-1, Y106G6D.4, SET-29, OAC-43, Y16B4A.2
Hydrolase activity	7	8	Y54F10BM.3, ATIC-1, MATH-33, PPH-4.1, C17H12.3, ABHD-11.2, Y16B4A.2, MRP-7
Lipid binding	4	2	CED-12, FER-1
Structural molecule activity	4	2	SYD-2, TTN-1
Cytoskeletal protein binding	3	1	TTN-1
Transferase activity	3	6	AC3.12, ATIC-1, SET-29, OAC-43, TTN-1, Y106G6D.4

Phosphoprotein phosphatase activity	2	3	Y54F10BM.3, PPH-4.1, C17H12.3
Peptidase activity	1	2	MATH-33, Y16B4A.2
Kinase activity	1	2	MATH-33, Y16B4A.2
Receptor activity	1	2	AC3.12, LGC-29
Receptor binding	1	1	SLT-1
Protein kinase activity	1	3	AC3.12, Y106GG6D.4, TTN-1
Signal transducer activity	1	1	AC3.12

GO Slim Categories in bold indicate $\alpha = 0.05$ significance for functional enrichment in the GoMiner output (see Methods). Genes in bold indicate their presence in any significantly enriched GO child category at $\alpha = 0.05$ for that particular GO Slim Category (p values not shown).